PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION International Bureau



(51) International Patent Classification 6:		11) International Publication Number:	WO 98/28971
A01K 67/027, 67/033, C12N 15/00	A2	(43) International Publication Date:	9 July 1998 (09.07.98
(21) International Application Number: PCT/L (22) International Filing Date: 31 December 1997 (30) Priority Data: 08/778,645 3 January 1997 (03.01.97) 60/080,222 3 January 1997 (03.01.97) (71) Applicant (for all designated States except US): STTY TECHNOLOGY CORPORATION [US 250, 3101 Iris Avenue, Boulder, CO 80301 (US) (72) Inventor; and (75) Inventor/Applicant (for US only): LINK, [US/US]: 1837 Spruce Street, Boulder, CO 803 (74) Agents: MONROY, Gladys, H. et al.; Morrison LLP, 755 Page Mill Road, Palo Alto, CA 94304	UNIVI (US]; Su (). Christop 02 (US).	BY, CA, CH, CN, CU, CZ, DE, GH, GM, GW, HU, ID, IL, IS, LC, LK, LR, LS, LT, LU, LV, MX, NO, NZ, PL, PT, RO, RU, TJ, TM, TR, TT, UA, UG, US, patent (GH, GM, KE, LS, MW, S patent (AM, AZ, BY, KG, KZ, M patent (AT, BE, CH, DE, DK, E LU, MC, NL, PT, SE), OAPI pa CM, GA, GN, ML, MR, NE, SN Published Without international search requipon receipt of that report.	DK. EE, ES, FI, GB, GF, JP, KE, KG, KP, KR, KZ, MD, MG, MK, MN, MW, SD, SE, SG, SI, SK, SI, UZ, VN, YU, ZW, ARIP SD, SZ, UG, ZW), Eurasit ID, RU, TJ, TM), Europea S, FI, FR, GB, GR, IE, 17 attent (BF, BJ, CF, CG, C4, TD, TG).

(54) Title: BETA-AMYLOID TOXICITY

(57) Abstract

The present invention relates to methods and compositions for visualization of the toxic effects of transgenes *in vivo*. In particular, the present invention provides methods and compositions for the production and use of transgenic, including dually transgenic, *Caenorhabditis* elegans for visualization of the toxic effects of β -amyloid accumulation in vivo.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

PCT/US97/23819 WO 98/28971

BETA-AMYLOID TOXICITY

FIELD OF THE INVENTION

The present invention is in the field of medicine and molecular biology. In particular, the invention provides transgenic non-human animals in which expression of a reporter gene is induced by a toxic transgene. The invention also provides methods and compositions for *in vivo* visualization of the toxicity associated with toxic proteins. In addition, the present invention provides methods and compositions for the production and use of transgenic non-human animals for screening candidate drugs to assess and treat toxicity.

10

15

20

25

5

BACKGROUND OF THE INVENTION

Investigation of the molecular mechanisms of various diseases has suggested that many pathologies are caused by expression of proteins that are immediately or gradually deleterious to the cells where they are expressed. Among these diseases are Alzheimer's disease, prion diseases, Huntington's disease, and amyotrophic lateral sclerosis. Several of these diseases are thought to result from aberrant folding of proteins, which results in the accumulation of toxic proteins or protein aggregates.

Alzheimer's disease is associated with the majority of dementia cases in the United States, with an estimated 2 million people afflicted with the disease, and a mortality rate of approximately 100,000 people per year (See, R.W.P. Cutler, "Degenerative and Hereditary Diseases." in Medicine, Scientific American, New York, (1988), pages 11 (IV):1-13; and R. Katzman (1986) N. Engl. J. Med. 314:964). It has been estimated that the total cost for nursing home care alone of Alzheimer's patients exceeds \$13 billion/year (See, M.M. Heckler (1985) Am. Psychol. 40:1240). According to the Centers for Disease Control (CDC), mortality due to Alzheimer's disease in the United States increased 10-fold between 1979 and 1987 (0.4 per 100.000 to 4.2 per 100.000) (See, "Reported death rate for Alzheimer's is up tenfold since 1979," (1990) Clin. Psychiatr. News 18:21).

Patients suffering from Alzheimer's disease typically suffer progressive memory deficit, progressive decline in cognitive functions, anxiety, depression, visuospatial and speech deficits, delusions, personality changes, motor skill deterioration, loss of verbal

ability, and incontinence. Eventually, patients are completely incapacitated and disoriented, requiring total care. The course of Alzheimer's disease ranges from less than three years to over 20 years before death occurs. However, in typical cases, it progresses at a fairly constant rate, with an average duration of 6 to 10 years.

5

10

15

Alzheimer's disease is not the only cause of dementia. Indeed, there are over fifty recognized causes of dementia. As some causes of dementia are amenable to treatment, differential diagnosis of patients suffering dementia is particularly important. Because at least 20% of clinically diagnosed patients were found at autopsy to have had conditions other than Alzheimer's disease, the National Institute of Neurological and Communicative Disorders and Stroke (NINCDS) and Alzheimer's Disease and Related Disorders

Association (ADRDA) retined the clinical diagnostic criteria for Alzheimer's disease (See, M.A. Jenike, "Psychiatry," in Medicine, Scientific American, New York, [1991], pages 13 (V):1-5). Based on these criteria, the diagnosis of Alzheimer's disease may be "definite" (i.e., requiring examination of brain tissue), "probable" (i.e., patients have deficits in two or more areas of cognition, insidious onset of disease, progressive worsening of memory and other cognitive functions, and normal consciousness levels), or "possible" (i.e., patients meet the criteria for probable Alzheimer's disease, but exhibit variations in the disease course or have a systemic illness that is sufficient to cause dementia, but is not considered

20

to be the cause of the dementia).

The numerous varieties of dementia, and variations in patient presentations, often make diagnosis problematic. Thus, the NINCDS/ADRDA criteria are very detailed, and necessitate the thorough examination of patients with suspected Alzheimer's disease. Currently, the only way to obtain a definite diagnosis is by post-mortem histological examination of brain tissue for the presence of senile plaques.

25

The histopathological lesions of Alzheimer's disease include neuritic or senile plaques, neurofibrillary degeneration, and granulovacuolar neuronal degeneration. The senile plaques usually contain a core of insoluble, amyloidic extracellular material (" β -amyloid) surrounded by a halo of neurofibrillary tangles and dystrophic neurons. The primary protein component of the amyloidic core of senile plaques is a 4.2 kd amyloid β peptide (often referred to as " $A\beta$ "). A number of other proteins have also been identified

30

as components of senile plaques, including anti-chymotrypsin and apolipoprotein E. The major component of the neurofibrillary tangles is an abnormally phosphorylated microtubule-associated protein referred to as "tau." The β-amyloid protein in senile plaques is a small glycoprotein (i.e., a 39-43 amino acid protein, derived from the larger amyloid precursor protein), that has been detected in non-neural tissues (e.g., skin, subcutaneous tissue, and intestines) and blood vessels of Alzheimer's disease patients (See, Cutler, supra). Deposits of amyloid may be detected by their ability to bind specific dyes, such as Congo red or thioflavin S. Thus, it has been further hypothesized that detection of this protein may serve as a potential diagnostic aid in the assessment of patients suffering from dementia.

5

10

15

20

25

30

Alzheimer's disease usually arises spontaneously, although genetics play a role in development of the disease. For example, the histopathologic lesions of Alzheimer's disease have been regularly observed in the brains of older patients dying of Down's syndrome (*See, Cutler, supra*). It is known that the β-amyloid gene resides on chromosome 21. As Down's patients have an extra copy of this chromosome, there is an increased expression of this chromosome in these patients. It has been hypothesized that increased expression of the β-amyloid gene may regulate the formation of amyloid plaques ("senile plaques") in these patients. In addition, four loci were recently identified as playing a role in the genetic susceptibility of Alzheimer's disease (*See, Pericak-Vance and Haines* (1995) *Trends Genet.* 11:504).

Other potential risk factors for the development of Alzheimer's disease include environmental factors (e.g., head trauma, smoking, and exposure to heavy metals), sociological factors (e.g., depression and educational level), biological factors (e.g., increasing age and hyperthyroidism), and a family history of Alzheimer's disease, Down's syndrome, or Parkinson's disease (See, Pericak-Vance and Haines, supra). Nonetheless, despite recent advances, the exact etiology and pathogenesis of Alzheimer's disease remain largely unknown.

In addition to the problems associated with diagnosing Alzheimer's disease, improvements are needed in the area of treatment. Many agents have been tested for their ability to treat the cognitive decline associated with Alzheimer's disease. For example,

PCT/US97/23819 WO 98/28971

various cholinergic enhancers (eg, choline and lecithin) have been tested. Unfortunately, cholinergic precursors have been shown to be not useful, although some drugs that stimulate cholinergic transmission may be helpful in some patients. One example is physostigmine, a compound that prevents the synaptic breakdown of acetylcholine. However, the overall clinical effect of this drug has not been as dramatic as initially hoped (See, M.A. Jenike, "Psychiatry," in Medicine, Scientific American, New York, (1991), pages 13 (V):1-5).

Other drugs, such as tetrahydroaminoacridine (THA or tacrine), a centrally acting anti-cholinesterase, have been tested. In a large multi-center trial of THA, liver enzyme abnormalities were reported and the preliminary results indicated that, at least at low dosages, THA is not an effective treatment of Alzheimer's disease (Jenike, supra). Ergoloid mesylates (Hydergine) is an extremely safe compound, and remains the most commonly prescribed drug for patients with Alzheimer's disease. However, the overall effects of the drug are at best minimal.

The mechanisms of toxicity due to accumulation of β -peptide are currently the subject of much investigation, and no definitive causes of toxicity have yet been established. (See. Benzi and Moretti (1995) Neurobiology of Aging, 16:661-674). Because of this uncertainty, efforts to develop mechanism-based treatment regimens have not been possible. Compounds presently in use to treat Alzheimer's disease only serve to alleviate the systemic effects associated with the disease.

In order to study disease mechanisms and genetic-based phenomena, animals in which a foreign gene has been inserted have been described by various researchers. International Patent Application WO 96/03034 describes insertion of retroviral vectors into fish, in order to produce fish with desirable traits or to study development. Various transgenic animal models for Alzheimer's disease are described in International Patent Applications WO 93/14200, WO 93/02189, WO 94/12627, WO 94/23049, and European Patent Publication EP 653154. Typically these transgenic animals are mice or other mammals; however, β -amyloid peptide has been expressed in the nematode Caenorhabditis elegans (Link (1995) Proc. Natl. Acad. Sci. 92: 9368). The use of C. elegans for investigation of mutant forms of the C. elegans genes mec-4 and deg-1, which

-4-

20

5

10

15

25

30

PCT/US97/23819 WO 98/28971

cause neurodegeneration, is described in U.S. Patent No. 5,196,333. Drawbacks to mammalian animals as model systems are the relatively long generation time, which makes mammals less desirable for high-throughput screening of potential pharmaceuticals, and the difficulty in studying the molecular processes of interest without sacrificing the animal for cell and tissue analysis.

Reporter genes are genes that encode proteins or other compounds that can be detected by a variety of methods, and which "report" the occurrence of successful introduction and expression of gene sequences. β-galactosidase and luciferase are examples of such reporter genes. Recently, the gene for green fluorescent protein (GFP) of the jellyfish *Aequorea victoria* has been described in U.S. Patent No. 5,491,084. This gene can provide a method for indicating expression of a gene via fluorescent detection of GFP. The drawback to using GFP in mammalian systems is as described above, that is, the need to sacrifice the mammal in order to analyze the cells and tissues of interest. This drawback is especially severe in the context of large-scale screening of potential therapeutic compounds, and U.S. Patent No. 5,491,084 does not describe a system suitable for high-throughput screening of pharmaceuticals for activity against toxic proteins expressed in cells.

Thus, methods and animal systems are needed to screen drugs quickly and inexpensively for their effects on proteins and other substances associated with Alzheimer's disease. Convenient methods and animal systems for screening drugs for other neurodegenerative diseases, such as prion diseases. Huntington's disease, and amyotrophic lateral sclerosis, are also desirable.

SUMMARY OF THE INVENTION

The present invention provides methods and compositions useful for the production and use of dually transgenic animals, in particular *Caenorhabditis elegans*. In addition, in one embodiment, the present invention provides transparent animals that express a reporter gene inducible by a toxic transgene. In a preferred embodiment, the toxic transgene encodes β -peptide. Although it is not intended that the present invention be limited to any particular reporter, in a preferred embodiment, the reporter gene is green fluorescence

- 5 -

30

5

10

15

20

25

protein (GFP). It is contemplated that the effects or product of the reporter gene be observable (*t.e.*) the reporter gene is expressed). For example, in embodiments in which GFP is the reporter, its presence is detected using fluorescence microscopy. In addition, in these embodiments, the animals do not need to be sacrificed in order to observe the expression of the reporter gene (*i.e.*) the detection may be accomplished on living animals). If other reporters are used, other detection methods may be necessary. For example, *lacZ* expression may be detected by exposing the tissues of the animal to the substrate for the gene (*i.e.*) β-galactoside), and observing for the presence of blue dye in the tissues. However, this method requires that the animal be sacrificed in order to observe the expression of the reporter gene.

5

10

15

20

25

In one alternate embodiment, the genome of the transparent animal further contains a heterologous marker gene. In a preferred embodiment, the marker gene is rol-6(su-1006). In the embodiments in which rol-6(su-1006), expression of the marker gene is accomplished by observing the living animals for the rolling behavior that results due to the expression of the abnormal collagen gene. In contrast, wild-type animals exhibit normal locomotion and do not exhibit this rolling behavior, as their collagen is normal.

In an alternative embodiment, the first transgene of the dually transgenic non-human animal comprises pCL25. In another alternative embodiment, the second transgene of the dually transgenic non-human animal comprises pCL12. In one alternate embodiment, the genome of the transparent animal further contains a heterologous marker gene. In a preferred embodiment, the marker gene is rol-6(su-1006). In the embodiments in which rol-6(su-1006), expression of the marker gene is accomplished by observing the living animals for the rolling behavior that results due to the expression of the abnormal collagen gene. In contrast, wild-type animals exhibit normal locomotion and do not exhibit this rolling behavior, as their collagen is normal.

In a preferred embodiment, the transparent animal is selected from the class Nematoda. Although it is not intended that the transparent animal of the present invention be limited to any specific animal, in a particularly preferred embodiment, the transparent animal is *Caenorhabditis elegans*.

PCT/US97/23819 WO 98/28971

5

10

15

20

25

30

In an alternate embodiment the cells of the transparent animal display toxicity resulting from the accumulation of β -peptide within the cells. In one preferred embodiment, the genome of the transparent animal comprises SEQ ID NO:5 and SEQ ID NO:8.

The present invention also provides methods for producing dually transgenic non-human animal comprising: providing: a first and second non-human animal; a first transgene comprising β -peptide: and a second transgene, comprising a reporter: introducing the first transgene into the genome of the first non-human animal to produce a first transgenic animal, and introducing the second transgene into the genome of the second non-human animal to produce a second transgenic animal; and mating the first transgenic animal with the second transgenic animal to produce a dually transgenic animal, wherein the β -peptide and the reporter are expressed.

In one embodiment of the methods of the present invention the dually transgenic non-human animal is transparent. In a preferred embodiment, the transparent animal is a nematode, while in a particularly preferred embodiment, the animal is *Caenorhabditis* clegans.

In an alternative embodiment, the first transgene of the dually transgenic non-human animal comprises pCL25. In another alternative embodiment, the second transgene of the dually transgenic non-human animal comprises pCL12. In one alternate embodiment, the genome of the transparent animal further contains a heterologous marker gene. In a preferred embodiment, the marker gene is rol-6(su-1006). In the embodiments in which rol-6(su-1006), expression of the marker gene is accomplished by observing the living animals for the rolling behavior that results due to the expression of the abnormal collagen gene. In contrast, wild-type animals exhibit normal locomotion and do not exhibit this rolling behavior, as their collagen is normal.

The present invention also comprises methods for testing compounds for anti-toxic effects, comprising: providing a dually transgenic non-human animal expressing a toxic transgene and a reporter; a composition comprising a test compound in a form suitable for administration such that the compound is bioavailable in the cells of the animal; and administering the test compound to the non-human animal. The toxic transgene induces

5

10

15

20

25

30

expression of the reporter gene. In a preferred embodiment, the toxic transgene is β -peptide and the reporter gene is GFP. In one embodiment, the method further comprises the step of measuring a reduction or increase in the expression of the reporter by the dually transgenic non-human animal and thereby identifying a compound as therapeutic. In a particularly preferred embodiment of the methods, the compounds inactivate the β -peptide expressed by the dually transgenic animal.

In one embodiment of the methods for testing compounds for β-peptide toxicity, the dually transgenic non-human animal is transparent. In a preferred embodiment, the transparent animal is a nematode, while in a particularly preferred embodiment, the animal is *Caenorhabditis elegans*. In an alternative embodiment, the first transgene of the dually transgenic non-human animal comprises pCL25. In another alternative embodiment, the second transgene of the dually transgenic non-human animal comprises pCL12. In one alternate embodiment, the genome of the transparent animal further contains a heterologous marker gene. In a preferred embodiment, the marker gene is *rol-6(su-1006)*, which produces effects described herein.

In one alternative embodiment, one or more test compounds are tested for their ability to counter the toxic effects of transgene product. It is also contemplated that the test compounds will be tested for their ability to prevent the expression of the toxic transgene, for example, β -peptide.

In alternative embodiments, dually transgenic animals in which the expression of, or effects of transgene toxicity, are reduced or eliminated by the test compounds, are mated. The progeny of these matings are also then tested for the effects of test compounds on the expression of, or effects of toxicity. In yet another embodiment, the progeny of these matings are used in other assay systems for the identification of therapeutic compounds.

In yet another alternative embodiment, the dually transgenic animals of the present invention are used to identify methods suitable for the diagnostic testing of Alzheimer's disease or pathology due to other diseases. Thus, the dually transgenic animals of the present invention are used to develop assays suitable for use in humans or animal models of Alzheimer's disease.

DESCRIPTION OF THE FIGURES

Figure 1 is a diagram illustrating the construction of dual transgenic animals expressing β -amyloid peptide and an hsp/GFP reporter.

Figure 2 shows the DNA sequence (SEQ ID NO:5) and restriction map for pCL12. Figure 3 is a graphic map of pCL12.

Figure 4 shows the DNA (SEQ ID NO:6), and amino acid sequence (SEQ ID NO:7) of pCL12 from nucleotide 1071 through 1253 (*i.e.*, the β -(1-42) nucleic acid and amino acid sequence).

Figure 5 shows the DNA sequence (SEQ ID NO:8) and restriction map for pGFP-TT.

Figure 6 is a graphic map of pGFP-TT.

5

15

20

25

30

Figure 7 shows the DNA sequence (SEQ ID NO:9) and restriction map for pCL25.

Figure 8 is a graphic map of pCL25.

Figure 9 shows the DNA sequence (SEQ ID NO:10) of rol-6.

Figure 10 shows the amino acid sequence (SEQ ID NO:11) of rol-6.

DESCRIPTION OF THE INVENTION

The present invention provides methods and compositions useful for the production and use of transgenic animals. The methods and transgenic animals of the invention also provide an efficient and effective system for screening drug effective in ameloriating the effects of toxic gene products.

The practice of the present invention will employ, unless otherwise indicated, conventional techniques of molecular biology, microbiology, cell biology and recombinant DNA, which are within the skill of the art. *See, e.g.*, Sambrook, Fritsch, and Maniatis, MOLECULAR CLONING: A LABORATORY MANUAL, 2nd edition (1989); CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, (F.M. Ausubel et al. eds., 1987); the series METHODS IN ENZYMOLOGY (Academic Press, Inc.); PCR 2: A PRACTICAL APPROACH (M.J. McPherson, B.D. Hames and G.R. Taylor eds., 1995); ANIMAL CELL CULTURE (R.I.

Freshney, Ed., 1987); and ANTIBODIES A LABORATORY MANUAL (Harlow et al. eds., 1987).

Animals

5

10

15

20

25

30

The methods and systems described herein can be practiced with any non-human animal. In a preferred embodiment, the animal used is transparent for at least part of its life cycle, for example, oocytes (*Xenopus* or others); larvae; pupae, fish (zebrafish or others); and nematodes, such as *C. elegans*. *C. elegans* is particularly preferred because it is transparent throughout its life and has been extensively studied as a model to determine developmentally regulated gene expression, as well as pattern formation. The visibility of all of the animals' cells throughout their life cycles, has led to the complete determination of the cell lineages, and a detailed description of the morphogenic changes that occur during *C. elegans* development. Rapid methods for introduction of cloned DNA into the *C. elegans* germ line has provided means to study molecular function and expression in vivo, as the effects of the gene expression are usually readily visible in the transparent worms.

Reporter Molecules

The methods and compositions of the present invention make use of reporter genes to monitor gene activity. Visible and quantifiable reporter genes are known and described in the art. Successfully used reporter molecules in gene fusion vectors in studies with *C. elegans* include the *E. coli lacZ* coding region (*See e.g.*, A. Fire *et al.*, (1990) *Gene* 93:189-198), and the *Aequora victoria gfp* coding region (*See e.g.*, M. Chalfie *et al.*, (1994) *Science* 263:802-805; D.C. Prasher *et al.*, (1992) *Gene* 111:229-233; and Genbank Accession #M62654), which produces an intrinsically fluorescent protein. Although various fusion expression vectors have been used and reported in the literature, problems have been encountered (*See e.g.*, A. Fire, "Fire Lab Vector Kit"--June 1995). For example, with *lacZ*, there have been many reports in which it was not possible to correlate transgene expression patterns with physiological expectations, or only weak correlations have been possible. Ectopic expression is often frequently seen with short promoters, and occurs

PCT/US97/23819 WO 98/28971

most prominently in the gut and pharynx of the affected animals. This may be due to weak promoter and/or enhancer signals in the vectors. Expression pattern deficits of *lacZ* fusions have been classified into three groups. In the first, transgenes are expressed in the correct tissue, but mosaic expression is observed (*i.e.*, only a subset of the cells stain during the detection methods). This has been observed even with integrated high-copy transgenes. In the second group, expression in a single tissue or cell population is not seen with the transgene. Finally, there has been the failure of transgene fusion constructs to show expression in the pre-12 cell embryo, or in any embryonic, larval, or adult germ line.

The *gfp* fusion vectors were developed as an alternative to the *lac*Z markers, but have been available for a much shorter time (*See e.g.* Chalfie *et al.*, (1994), *supra*; U.S. Patent No. 5.491.084). Initial reports indicate that the fluorescence pattern appears to be more restricted than that exhibited by equivalent *lac*Z fusion. In addition, there are still problems associated with the germline expression of the transgenes and ectopic expression. Indeed, in some cases, the problems with ectopic expression have been exacerbated by these "improvements."

The present invention overcomes many of these problems described in the art. The transparent animals and dually transgenic animals described herein provide easily visible, stably expressed systems in which the toxic transgene and reporter gene are expressed similarly to endogenous chromosomal gene expression. These animals provide an excellent system for screening compounds having effects on the toxic transgene. The present invention, therefore, provides model systems for the study of human diseases and methods of identifying therapeutic compounds using these animal systems.

Inducible Promoters

5

10

15

20

25

30

As discussed above, the reporter genes are operably linked to an inducible promoter. The promoter is induced by the toxic transgene, for example by the gene product of the toxic transgene. It is also contemplated that the toxic transgene can act to induce the promoter indirectly, for example by disrupting other cellular proteins or functions. Suitable inducible promoters are available and can be readily determined by those skilled in the art. Non-limiting examples of promoters which are induced by

"stresses" include the metallothionein gene promoter (*mtl*-1 or *mtl*-2; *e.g.*, Genbank Accession #M92910, #M11794, #X00504, and #X00953), and the *C. elegans* amyloid precursor protein (APP) homolog *apl*-1 gene promoter (*See e.g.*, Daigle and Li, Proc. Natl. Acad., Sci., 90:12045-12049 [1993]). Other potential promoters include those from other inducible heat-shock genes; at least one of the known *C. elegans hsp*70 genes is strongly heat-inducible. It is also contemplated that promoters from genes known to be upregulated under stress conditions in other systems (*e.g.*, superoxide dismutase, catalase, glutathione reductase, etc.) may also be useful.

5

10

15

20

25

In one embodiment, as described in the Examples below, a strong muscle-specific promoter was used to express a potentially secretable form of the β -peptide, so as to generate significant extracellular levels of β -peptide. This was accomplished in order to mimic the situation that may exist in the human brain, and allow observations of cell-external neurotoxicity. The present inventor has also demonstrated that the promoter must be chosen so that enough β -peptide is produced to cause physiological effects, but not so much to kill the animals.

The unc-54/ β -(1-42) minigene was constructed with a modified signal sequence that has been previously shown to allow secretion of a her-1 protein product that is ectopically expressed in muscle cells (M.D. Perry $et\ al.$, (1993) $Genes\ Dev.$, 7:216-228). The development of C. elegans transgenic for expression of β -amyloid was previously described by the inventor (Link, (1995) $Proc.\ Natl.\ Acad.\ Sci.,\ 92:9368-9572$).

However, in early experiments, it was observed that β -peptide deposits were not convincingly detected outside of the muscle cells, when tested with the antibodies described in Example 4, below. While an understanding of the mechanism is not necessary for the practice of the present invention, it is apparent that the majority of β -peptide expressed by the transgenic worms is retained in the muscle cells and is responsible for the pathology observed in the muscle cells.

It is also contemplated that β -peptide expression will be directed to other tissues. through utilization of appropriate promoters. For example, it is contemplated that animals expressing β -peptide in the intestine may be particularly useful to analyze compounds such

as drugs for their effect on β-peptide, as these cells readily take up exogenous compounds administered orally.

Transgenes

5 The transgene element of the present invention can be any sequence which is able to induce the promoter operably linked to the reporter gene. As noted, the transgene is "toxic" in the sense that it disrupts cellular function in some way. Preferably, the transgene encodes a protein that is toxic to the host cells and/or organism in that it causes deleterious effects to the host, for instance interfering with the hosts ability to survive and/or grow. The toxic transgenes encode proteins that disrupt cellular function directly 10 (e.g., the gene product is toxic) or indirectly (e.g., the sequence of the transgene disrupts cellular function by some mechanism other than its gene product). Toxic proteins are distinguishable from simple chemical toxins (e.g., heavy metals and the like) by their

antigenicity and higher molecular weight.

' In one embodiment, the toxic transgene comprises a gene encoding for an amyloidic protein, for example beta amyloid peptide, prion protein variants, transthyretin variants, gelsolin variants, cystatin variants, lysozyme variants and the like. In another embodiment, the transgene encodes a protein containing polyglutamine resulting from triplet-repeat expansion such as huntingtin (a protein that has been implicated in Huntington's Chorea), ataxin-1 or ataxin-2. Alternatively, proteins associated with inherent amyelotopic lateral sclerosis (ALS) for example, superoxide dismutase 1 variants and over-expressed neurofilament protein, can be used. It will be understood that the transgene can encode for an entire toxic protein or, alternatively, a functional (i.e. toxic)

25

30

20

15

Transgenic Animals

In one aspect, the present invention includes a dually transgenic non-human animal. In one embodiment, the development of these dually transgenic animals involves the production of two C. elegans lines with a single transgene in each line, designated as CL2005 and CL2070. C. elegans was chosen due to the observation that the toxic effect of

β-amyloid has been reported to occur in these animals, in addition to humans. Line CL2005 exhibited muscle-specific expression of human β peptide, while line CL2070 exhibited stress-inducible expression of GFP. GFP was incorporated into the present invention as it provides a marker (i.e., "reporter") that is readily visible in living worms. Thus, dually transgenic animals both express β-peptide in their muscle cells and under appropriate conditions for observation (i.e., fluorescence microscopy), exhibit green fluorescence in these cells.

Mating of these lines resulted in the production of dually transgenic animals, in which the presence of β -amyloid was easy to detect, due to the fact that the dually transgenic animals expressing β -amyloid glowed green. Thus, the present invention provides an easily detectable method for the specific expression of β -peptide in dually transgenic animals. This ease of detection provides great advantages for the development of methods to analyze the effects of β -amyloid *in vivo*. In addition, because the dually transgenic animals are not killed in order to detect the presence of β -amyloid, as would be required if other reporters (*e.g., lacZ* or luciferase) were used, the animals may be useful for screening compounds for their effects on β -amyloid in an *in vivo* situation. Thus, the use of an reporter which is detectable *in vivo* provides significant advantages over currently available methods.

Other Advantages of Transgenic C. elegans

5

10

15

20

25

30

The present invention also includes methods of screening compounds for their ability to prevent or inhibit toxicity due to the toxic transgene. In these drug screening embodiments, the methods and animals described above can be used to identify potential therapies. For example, in assays to develop compounds to block the expression of β -amyloid or counter its toxic effects, transgenic worms (e.g., green fluorescing worms with muscle-specific β -peptide expression), would be observed for their loss of fluorescence. In assays where the compound is effective in countering the expression of β -peptide or effects of β -peptide, the animals would no longer glow green. It is easy to visually screen and quantify (using commercially available equipment) for animals that have lost the ability to glow, and then use them to detail the interaction between the compound and the toxic

transgene. The present invention provides a fast and efficient screening system, for PCT/US97/23819 example by using commercially available equipment for assaying multiple compounds at

Thus, the transgenic animals of the present invention expressing a detectable reporter gene triggered by a toxic transgene provide animal models for human diseases. For instance, a reporter gene (such as GFP) induced by expression of β-peptide provides an animal model for human Alzheimer's disease. It is also not intended that the present invention be limited to animals expressing β -peptide in conjunction with GFP. It is contemplated that expression of other proteins associated with dementia and/or 10

Alzheimer's or other diseases would also find use in conjunction with GFP in C. elegans as well. For example, it is contemplated that proteins (including mutated versions of proteins) such as the presentlins associated with Alzheimer's disease (See e.g., J. Marx (1996) Science 274:1838-1840) will be used in conjunction with GFP and full-length APP in dually transgenic Celegans. Other toxic proteins which provide animal models for 15 other human diseases are described herein.

These transgenic animals also provide a means to screen compounds for their ability to decrease or eliminate the toxicity. Examples of compounds suitable for testing using the transgenic animals of the present invention include, but are not limited to. Congo Red. tumor necrosis factor (TNF), estrogen, tacrine (9-amino-1.2,3,4-tetrahydroacridine). dihydroepiandrosterone (DHEA), compounds that inhibit ApoE4, and others, commercially available from suppliers such as Sigma. Compounds to be tested for antitoxic activity are administered to the same number of dually transgenic animals (e.g., generated using the methods described in the Examples. below) from the control group and the treatment group, and the presence or absence of reporter gene used as a measure of efficacy.

20

25

30

_982897147 .

The compounds being tested can be administered using any suitable route (e.g., oral, parenteral, controlled-release transdermal methods, and implants, etc.). In one preferred route, the compounds to be tested are suspended in the growth media provided to the worms. Generally speaking, the route of administration will depend on the stability of the compound, the susceptibility of the compound to "first pass" metabolism, the

concentration needed to achieve a therapeutic effect, and the like. Following initial screening, a compound that appears promising (i.e.), which increases the number of worms which display reduced β -peptide toxicity) is further evaluated by administering various concentrations of the compound to additional transgenic animals in order to determine an approximate therapeutic dosing range.

5

10

15

20

25

30

Another screening method involves the crossing of the transgenic worms of the present invention with other transgenic worms. The animals are observed after treatment, in the presence and absence of the test compound(s), with the effects on the toxic transgene being gauged either by crude survival or the presence/absence of the reporter gene. It is also contemplated that the methods of the present invention be modified so as to provide means to analyze disease-related proteins believed to have dominant toxic effects, by substituting appropriate sequences for the β-peptide sequences used as described in the Examples. For example, proteins such as transthyretin (known to be associated with familial amyloid polyneuropathy; See e.g., Christmason et al., (1991) FEBS 281:177-180; Genbank Accession #D00096), and variant superoxide dismutase (known to be associated with familial amyotrophic lateral sclerosis [Lou Gehrig's disease]), prion proteins, A4 amyloid protein (See e.g., Ponte et al., (1988) Nature 331:525-527; Salbaum et al., U.S. Patent No. 5,151,508, herein incorporated by reference), APP (See e.g., Kitaguchi et al., (1988) Nature 331:530-532; Sata et al., EP Appln. 94117512.7; Scott et al., WO 9412627; Wadsworth et al., WO 9314200; Gearhart et al., WO 9423049; and Neve et al., WO 9302189, all of which are herein incorporated by reference), other amyloidic proteins (e.g., variant lysozymes and amylin peptide), and other proteins associated with neuronal degeneration (See e.g., U.S. Patent No. 5,196,333), will be used in the methods of the present invention by substituting the appropriate nucleic acid sequences encoding the protein of interest for the β -peptide described in Example 1. It is also contemplated that animals transgenic for dual proteins (e.g., β-peptide in combination with another protein. such as transthyretin) will be used. Thus, the present invention provides the means to analyze the effects of numerous genes and proteins in vivo.

Furthermore, the transgenic animals of the present invention provide distinct advantages over other transgenic animals currently used to analyze diseases such as

Alzheimer's disease. Due to the short gestation period of *C. elegans*, transgente animals can be produced much more rapidly than when mammals, such as mice are used. For example, transgenic mice overexpressing a 695 amino acid isoform of β-amyloid precursor did not show learning and memory impairment until they were 9-10 months of age (Hsiao et al., Science 274:99-102 [1996]). In contrast, the transgenic animals of the present invention express β-peptide and may be manipulated at a very early age, even in embryonic stages.

5

0

In addition, because the transgenic animals are easy to select based on the presence of the green fluorescence, invasive procedures such as surgery, necessary to analyze the effects of the toxic transgene in other animals (e.g., mice) are avoided.

Finally, the apparent toxicity of β-peptide in transgenic animals was found to be temperature-dependent. Transgenic animals maintained at 25.5°C were significantly more sick than those maintained at 16°C. The animals maintained at the higher temperature became paralyzed more quickly, failed to eat or grow, failed to show normal egg-laying, and many died before reaching adulthood. However, wild-type animals grow well at either of these temperatures. The CL2005 parental line was temperature-sensitive for viability, as it was not possible to propagate this line at the elevated temperature. In the CL2070 parental line, 25.5°C maintenance was not sufficient to induce the hsp/GFP construct. The hsp/GFP response was similarly temperature-dependent. It was difficult to detect GFP when the dual transgenic animals were raised at 16°C, but GFP induction was dramatic when the animals were raised at 25.5°C

These temperature dependency observations were exploited by the propagation of animals at 16° C, and then upshifting them to 25.5° C, when GFP induction was desired. This effect also has potential benefits in the analysis of compounds, as the animals can be pre-incubated in the presence of drugs or other compounds at the lower temperature before the upshift to the higher temperature, in order to ensure that the presumed protective effect of the drug was in place, prior to the strong induction of β -peptide toxicity. Importantly, these assays may be completed within one day and are suitable for rapid methods (e.g., the use of a microtiter format and a plate fluorimeter), so that literally thousands of compounds may be tested simultaneously.

5

10

15

20

25

30

In sum, the present invention provides methods and compositions useful as animal models for disease, as well as providing methods and compositions for disease therapy and prevention. The animal testing may be supplemented and confirmed by testing on human subjects. However, the transgenic animals of the present invention allow the testing of a large number of compounds, both various methods, including those known in the art.

DEFINITIONS

To facilitate understanding of the invention, a number of terms are defined below. The term "non-human animal" includes vertebrates such as rodents, arthropods, insects (e.g., Diptera), fish (e.g., zebrafish), non-human primates, ovines, bovines, ruminants, lagomorphs, porcines, caprines, equines, canines, felines, aves, etc. The non-human animal can be used at any stage in its development, for example oocyte, fetal, larval, pupal stages or the like. Preferred non-human animals are those that are transparent, such as certain nematodes, oocytes, larvae and fish.

As used herein, the term "transparent" is used in references to animals through which light will be transmitted. However, it is not intended that the amount of light transmittance be limited to any particular amount. For example, an animal is transparent so long as at least some light may be transmited through its body. Non-limiting examples of transparent animals are larval stages of some animals (e.g., flies) and oocytes (e.g., Xenopus oocytes). Particularly preferred non-human animals are selected from the nematodes (i.e., any animal in the Class Nematoda), most preferably Caenorhabditis elegans. However, it is contemplated that other transparent animals, such as zebrafish will be useful in the present invention. For example, transgenic zebrafish have been produced (See e.g., Lin et al., PCT Publ. WO9603034; incorporated herein by reference).

The "non-human animals having a genetically engineered genotype" of the invention are preferably produced by experimental manipulation of the genome of the germline of the non-human animal. These genetically engineered non-human animals may be produced by several methods including the introduction of a "transgene" comprising nucleic acid (usually DNA) into an embryonal target cell or integration into a chromosome of the somatic and/or germ line cells of a non-human animal by way of human

WO 98/28971

intervention, such as by the methods described herein. The process by which a DNA molecule becomes stably incorporated into another genome is referred to as "stable integration." Non-human animals which contain a transgene are referred to as "transgenic animal is an animal whose genome has been altered by the introduction of one or more transgenes.

5

10

15

20

25

30

The term "transgene" as used herein refers to a foreign gene that is placed into an organism by introducing the foreign gene into gonadal cells, embryonic cells, newly gene sequence) which is introduced into the genome of an animal by experimental introduced gene contains some modification relative to the naturally-occurring gene. A "toxic transgene" refers to a foreign gene which disrupts cellular function in some way. For example, the toxic transgene may produce a gene product (i.e., protein) which is toxic to the cell or organism. Alternatively, the transgene may disrupt other cellular proteins, or act as a DNA binding site without producing a protein

As used herein, the term "toxicity" refers to the production of toxic effects by any compound or substance. For example, while it is not necessary to the understanding and use of the present invention, β -peptide toxicity may arise due to the accumulation of β -peptide molecules. There are numerous manifestations of toxicity that may occur. For example, the toxicity associated with β -peptide may be manifested as negative effects on muscle cells of dually transgenic animals.

The terms "promoter element," "promoter," or "promoter sequence" as used herein, refer to a DNA sequence that is located at the 5' end of (i.e., precedes) a gene in a DNA polymer and provides a site for initiation of the transcription of the gene into mRNA. An instance binding of a transcription factor, stress, heat or the like

The term "reporter gene" as used herein refers to genes that encode proteins or other compounds that can be detected by a variety of methods. These genes "report" the occurrence of successful introduction and expression of sequences such as transgenes.

Non-limiting examples of reporter genes include antibiotic resistance genes, genes encoding enzymes and genes encoding other detectable proteins. Expression of the reporter gene is detected using methods known in the art. In a preferred embodiment, the reporter gene is a GFP gene, although lacZ β -galactosidase gene, or any other reporter system may be used to detect the successful production of transgenic animals. In particularly preferred embodiments, the reporter is a compound or protein which may be present or expressed within living animals. That is, it is not necessary to sacrifice the animal in order to detect the presence of the reporter. The type of the reporter gene system used is not critical to the invention, and it is contemplated that any system suitable for use with the transgenic animals of the present invention will be used.

5

10

15

20

25

30

As used herein, the terms "dual trangenic" and "dually transgenic" refer to animals or cells in which more than one transgene have been introduced. For example, the term is used in reference to cells which contain the sequences encoding β -peptide, and sequences encoding a reporter (e.g., GFP). However, it is not intended that the number of transgenes in the dually transgenic animals of the present invention be limited to two. For example, the transgenic animals may also contain another one or more marker genes (e.g., rol-6 sequences), in addition to β -peptide and reporter sequences. In the embodiments in which rol-6(su-1006), expression of the marker gene is accomplished by observing the living animals for the rolling behavior that results due to the expression of the abnormal collagen gene. In contrast, wild-type animals exhibit normal locomotion and do not exhibit this rolling behavior, as their collagen is normal.

The transgenic animals of the present invention are preferentially generated by introduction of the targeting vectors into gonad cells. Transgenes can be efficiently introduced into the cells by DNA transfection using a variety of methods known to the art, including electroporation, calcium phosphate co-precipitation, protoplast or spheroplast fusion, lipofection, and DEAE-dextran-mediated transfection. Transgenes may also be introduced into cells by retrovirus-mediated transduction or by micro-injection. In one preferred embodiment, the transgenes are injected into gonads of *C. elegans* as described by Mello *et al.* (1991) *EMBO J.*. **10**:3959-3970. Alternative methods for the generation of transgenic animals containing an altered gene are known to the art. For example,

embryonal cells at various developmental stages can be used to introduce transgenes for the production of transgenic animals. Different methods are used depending on the stage of development of the embryonal cell.

5

10

15

)

A wide variety of labels and conjugation techniques are known by those skilled in the art and can be used in various nucleic acid and amino acid assays. Means for producing labeled hybridization or PCR probes for detecting related sequences include oligolabeling, nick translation, and end-labeling or PCR amplification using a labeled nucleotide. Suitable reporter molecules or labels include those radionuclides, enzymes, fluorescent, chemiluminescent, or chromogenic agents as well as substrates, cofactors, inhibitors, magnetic particles and the like. The term may also be used in reference to proteins. For example, a variety of protocols which employ polyclonal or monoclonal antibodies specific for the β-peptide protein product are known in the art (See, the Examples). These antibodies can be used as markers for the expression of proteins. Examples include enzyme-linked immunosorbent assay (ELISA), radioimmunoassay (RIA) and fluorescent activated cell sorting (FACS).

The term "gene" refers to a DNA sequence that comprises coding sequences necessary for the production of a polypeptide or precursor. In some instances, a gene can also include control sequences. The polypeptide can be encoded by a full length coding sequence or by any portion of the coding sequence so long as the desired enzymatic activity is retained.

The term "gene of interest" refers to any gene, the manipulation of which may be deemed desirable for any reason, by one of ordinary skill in the art.

The term "wild-type" refers to a gene or gene product which has the characteristics of that gene or gene product when isolated from a naturally occurring source. A wild-type gene is that which is most frequently observed in a population and is thus arbitrarily designed the "normal" or "wild-type" form of the gene. In contrast, the term "modified" or "mutant" refers to a gene or gene product which displays modifications in sequence and or functional properties (*i.e.*, altered characteristics) when compared to the wild-type gene or gene product. It is noted that naturally-occurring mutants can be isolated; these are

tified by the fact that they have altered characteristics when compared to the wild-type

The terms "targeting vector" or "targeting construct" refer to oligonucleotide inces comprising a gene of interest flanked on either side by regulatory sequences. ably, the targeting vector is capable of homologous recombination such that the gene

As used herein, the terms "vector" and "vehicle" are used interchangeably in ice to nucleic acid molecules that transfer DNA segment(s) from one cell to another. The terms "expression vector" or "expression cassette" as used herein, refers to a mant DNA molecule containing a desired coding sequence and appropriate nucleic juences necessary for the expression of the operably linked coding sequence in a ir host organism. Nucleic acid sequences necessary for expression in prokaryotes nclude a promoter, an operator (optional), and a ribosome binding site, often along r sequences. Eukaryotic cells are known to utilize promoters, enhancers, and on and polyadenylation signals.

ne terms "in operable combination," "in operable order," and "operably linked" as n, refer to the linkage of nucleic acid sequences in such a manner that a nucleic sule capable of directing the transcription of a given gene and/or the synthesis of rotein molecule is produced. The term also refers to the linkage of amino acid n such a manner so that a functional protein is produced.

term "tissue-type specific" as it applies to a promoter, refers to a promoter that f selectively directing expression of a gene in a specific tissue. Similarly, the ne-specific promoter in the method of the present invention does not require cificity. In general, the requisite specificity is found where a plurality (or, bly, a majority) of cells in one tissue type express a gene of interest, while e.g., greater than 80%, and preferably greater than 90%, and more preferably 5%), of the cells in other tissue types do not. In one embodiment of the ion, the strong muscle promoter *unc*54 was used (See, Example 1). d herein the term "portion" when in reference to a gene refers to fragments of fragments may range in size from a few nucleotides to the entire gene

sequence minus one nucleotide. Thus, "an oligonucleotide comprising at least a portion of a gene" may comprise fragments of the gene or the entire gene.

As used herein, the term "substantially purified" refers to molecules, either nucleic or amino acid sequences, that are removed from their natural environment, isolated or separated, and are at least 60% free, preferably 75% free, and most preferably 90% free from other components with which they are naturally associated. An "isolated polynucleotide" is therefore a substantially purified polynucleotide

5

15

20

25

As used herein, the term "polymerase chain reaction" ("PCR") refers to the method of K.B. Mullis disclosed in U.S. Patent Nos. 4,683,195, 4,683,202 and 4,965,188, all of which are hereby incorporated by reference, which describe a method for increasing the 10 concentration of a segment of a target sequence in a mixture of genomic DNA without cloning or purification. This process for amplifying the target sequence consists of introducing a large excess of two oligonucleotide primers to the DNA mixture containing the desired target sequence, followed by a precise sequence of thermal cycling in the presence of a DNA polymerase. The two primers are complementary to their respective strands of the double stranded target sequence. To effect amplification, the mixture is denatured and the primers then annealed to their complementary sequences within the target molecule. Following annealing, the primers are extended with a polymerase so as to form a new pair of complementary strands. The steps of denaturation, primer annealing and polymerase extension can be repeated many times (i.e., denaturation, annealing and extension constitute one "cycle": there can be numerous "cycles") to obtain a high concentration of an amplified segment of the desired target sequence. The length of the amplified segment of the desired target sequence is determined by the relative positions of the primers with respect to each other, and therefore, this length is a controllable parameter. By virtue of the repeating aspect of the process, the method is referred to as the "polymerase chain reaction" (hereinafter "PCR"). Because the desired amplified segments of the target sequence become the predominant sequences (in terms of concentration) in the mixture, they are said to be "PCR amplified."

With PCR, it is possible to amplify a single copy of a specific target sequence in genomic DNA to a level detectable by several different methodologies (e.g., hybridization

with a labeled probe; incorporation of biotinylated primers followed by avidin-enzyme conjugate detection; and/or incorporation of ¹²P-labeled deoxynucleotide triphosphates, such as dCTP or dATP, into the amplified segment). In addition to genomic DNA, any oligonucleotide sequence can be amplified with the appropriate set of primer molecules. In particular, the amplified segments created by the PCR process itself are, themselves, efficient templates for subsequent PCR amplifications. Amplified target sequences may be used to obtain segments of DNA (*e.g.*, genes) for the construction of targeting vectors, transgenes, etc.

5

10

15

20

25

30

As used herein, the terms "PCR product" and "amplification product" refer to the resultant mixture of compounds after two or more cycles of the PCR steps of denaturation, annealing and extension are complete. These terms encompass the case where there has been amplification of one or more segments of one or more target sequences.

The term "Central Nervous System" refers to the "spinal cord" and the "brain." The spinal cord comprises white areas and grey areas. The grey area contains nerve cell bodies, whereas the white area is essentially comprised of myelinated nerves. The brain, which is also known as the encephalon, is that portion of the cerebrospinal axis which is contained in the cavity of the cranium. The brain comprises the two cerebral hemispheres, the interbrain, the mid-brain, the pons Varolli and cerebellum, and the medulla oblongata. The two hemispheres together with the parts derived from the thalamencephalon form the forebrain. The two cerebral hemispheres are separated by the longitudinal fissure and also comprise the bylvian fissure, the fissure of Rolando, and the parieto-occipital fissure. The lobes on the external surface of the brain comprise the frontal lobe, the parietal lobe, the occipital lobe, and the temporal lobe. Placed along the middle line of the brain are, among others, the rostrum and peduncles of corpus callosum, lamina cinera, optic commissure and the pituitary body. On each side of the middle line lies the frontal lobe, olfactory lobe and the hemisphere of cerebellum.

The terms "neuron." "neural cell," and "nerve cell" are used interchangeably to refer to a cell which is located in the nervous system. Nerve cells are composed of the nerve cell body (perikaryon), one or more dendrites, and an axon. Neurons can be classified according to the number of processes originating from the cell body. Thus,

unipofar neurons have a single process, bipolar neurons have one axon and one dendrite, while multipolar neurons (which are the most common) comprise more than two processes. The term "neuron" comprises cholinergic neurons and sensory neurons. As used herein, the term "cholinergic neuron" means a neuron in the Central Nervous System (CNS) and in the Peripheral Nervous System (PNS) whose neurotransmitter is acetylcholine. As used herein, the term "sensory neuron" includes a neuron which is responsive to environmental cues (e.g., temperature and movement) from, for example, the skin, muscle and joints of a mammal.

5

[()

15

20

25

30

The term "nerve" refers to two or more neurons arranged in linear sequence such that the axon of one neuron establishes a structural and functional link with the dendrite of a second neuron to form a "synapse."

The term "compound" refers to any chemical entity, pharmaceutical, drug, and the like that can be used to treat or prevent a disease, illness, sickness, or disorder of bodily function. Compounds comprise both known and potential therapeutic compounds. A compound can be determined to be therapeutic by testing using the testing methods of the present invention (i.e., a "test compound"). A "known therapeutic compound" refers to a therapeutic compound that has been shown (e.g., through animal trials or prior experience with administration to humans) to be effective in such treatment or prevention. In other words, a known therapeutic compound is not limited to a compound efficacious in the treatment or prevention of neural related disorders.

A compound is said to be "in a form suitable for administration such that the compound is bioavailable in the blood of the animal" when the compound may be administered to an animal by any desired route (e.g., oral, intravenous, subcutaneous, intramuscular, etc.) and the compound or its active metabolites appears in the blood of the animal in an active form. Administration of a compound to a pregnant animal may result in delivery of bioavailable compound to the embryonic progeny of the animal.

The "wild-type β -amyloid" or " β -peptide" gene and gene product refers to the nucleotide and amino acid sequences provided in SEQ ID NOS:6 and 7, respectively. Those skilled in the art will be well aware that certain modifications of SEQ ID NOS:6 and 7 can be made which will not interfere with the production of a polypeptide having an

_ - 25 -

activity indistinguishable from that of the wild-type β -amyloid; the present invention specifically contemplates these variant forms of β -amyloid. A "variant" of the β -peptide is defined as an amino acid sequence that differs by one or more amino acids. The variant may have "conservative" changes, wherein a substituted amino acid has similar structural or chemical properties (e.g., replacement of leucine with isoleucine). More rarely, a variant may have "nonconservative" changes, e.g., replacement of a glycine with a tryptophan. Similar minor variations may also include amino acid deletions or insertions (i.e., additions), or both. Guidance in determining which and how many amino acid residues may be substituted, inserted or deleted without abolishing biological or immunological activity may be found using computer programs well known in the art, for example, DNAStar software.

The term "an oligonucleotide sequence comprising at least a portion of a human β -amyloid gene" refers to a polynucleotide sequence (*i.e.*, a nucleic acid sequence) containing a nucleotide sequence derived from a human β -amyloid gene. This sequence may encode a portion or all of the β -amyloid protein; alternatively, this sequence may contain non-coding regions derived from the β -amyloid gene or a combination of coding and non-coding regions. The oligonucleotide may be RNA or DNA and may be of genomic or synthetic origin.

As used herein the term "portion" when in reference to a gene refers to fragments of that gene. The fragments may range in size from 10 nucleotides to the entire gene sequence minus one nucleotide. Thus, "an oligonucleotide comprising at least a portion of a gene" may comprise fragments of the gene or the entire gene.

An animal whose genome comprises a "heterologous marker gene" is an animal whose genome contains a marker gene not naturally found in the animal's genome. In one preferred embodiment of the present invention, the heterologous marker gene is a mutant collagen gene, such as the rol-6(su-1006) gene. However, it is intended that other marker genes will be used with success in the present invention, including other mutant collagen genes, as well as other marker genes commonly known to those in the art.

As used herein, the term "diagnostic assay" refers to methods for the diagnosis of disease, illness, and/or pathology. It is intended that the term encompass any methods for

diagnosis, including, but not limited to assays based on immunoreactivity (e.g.) radioimmunoassays, fluorescence immunoassays, enzyme immunoassays), histochemistry, dye retention or binding (e.g.) fixing of dyes such as Congo Red), nucleic acid based diagnostic methods (e.g.) identification of nucleic acid sequences associated with disease or pathology), etc.

EXPERIMENTAL

The following examples are provided in order to demonstrate and further illustrate certain preferred embodiments and aspects of the present invention and are not to be construed as limiting the scope thereof.

In the experimental disclosure which follows, the following abbreviations apply: eq (equivalents); M (Molar); μM (micromolar); N (Normal); mol (moles); mmol (millimoles); µmol (micromoles); nmol (nanomoles); g (grams); mg (milligrams); µg (micrograms); ng (nanograms); l or L (liters); ml (milliliters); µl (microliters); cm (centimeters); mm (millimeters); µm (micrometers); nm (nanometers); °C (degrees Centigrade); U (units), mU (milliunits); min. (minutes); sec. (seconds); % (percent); kb (kilobase); bp (base pair): PCR (polymerase chain reaction); β -(1-42) amino acids 1-42 of β-amyloid peptide; Tris (tris(hydroxymethyl)-aminomethane); BSA (bovine serum albumin); Fisher (Fisher Scientific, Pittsburgh, PA); Sigma (Sigma Chemical Co., St. Louis, MO.); Promega (Promega, Corp., Madison, WI); Perkin-Elmer (Perkin-Elmer/Applied Biosystems, Foster City, CA); Senetek (Senetek, PLC, Maryland Heights. MO): Boehringer Mannheim (Boehringer Mannheim, Corp., Indianapolis, IN); Stratagene (Stratagene, Inc., La Jolla, CA); and NEB (New England Biolabs, Beverly, MA). Unless otherwise indicated, the restriction enzymes used in these Examples were obtained from NEB. C. elegans is available from the Caenorhabditis Genetics Center. at the University of Minnesota, St. Paul, MN.

EXAMPLE 1

Assembly of the β -(1-42) Minigene

30

5

10

15

20

25

1-42) minigene used in subsequent experiments (See, ic C elegans engineered to express amyloidic human escribed below, these animals contained constructs in 54 promoter enhancer of C elegans drove the expression of lerived from human β-amyloid cDNA clones, as described ıman β-amyloid peptide in transgenic Caenorhabditis , 92:9368-9372 [1995]). ng the 42 amino acid β-amyloid peptide derived from n cDNA was assembled. The artificial signal peptide 52.81 was amplified under standard conditions, using "GGCCAAAGGACCC-3")(SEQ ID NO:1), and "SP-down" CCAGCAAGAT-3')(SEQ ID NO: 2), cleaved with Nhe1 nd inserted between the unique Nhel and Kpnl sites of construct "pCL2." This process resulted in a ree, such that the signal peptide cleavage site, as predicted by von Heijne, Nucl. Acids Res., 14:4683-4690 [1986]). Gly-Thr dipeptide encoded by the Kpnl site. ding amino acids 1-42 of β-amyloid, and which contained olified using standard methods, from human \(\beta\)-amyloid 4T4B (P. Ponte et al., Nature 331:525-527 [1988]), by GGGGTACCGATGCAGAATTCCGACATGA-3') (SEQ S'-CCCGAGCTCACGCTATGACAACACCGCCAAcation product was cleaved with KpnI and SacI, and I and Sac I sites of pCL2, to generate "pCL3." 12) minigene fragment was removed from this plasmid by d inserted between the unique Nhel and Sacl sites of The sequence of the β -(1-42) minigene was confirmed by coding strand only, by techniques known in the art. The 1:5), is shown in Figure 2. A graphic map of pCL12 is nows the DNA and amino acid sequences (SEQ ID NOS:

6 and 7, respectively) of pCL12 from nucleotide 1071 through 1253 (i.e., the β (1-42) nucleic acid and amino acid sequences) present in the construct.

EXAMPLE 2

Construction of the hsp/GFP Reporter

imple, the expression vector ("pCL25") containing the hsp/GFP reporter expression of the GFP-TT gene was constructed. This vector was used in ments (See. Example 3) to produce transgenic C. elegans was '-TT is a modified form of GFP that contains Ser65Thr and Ile167Thr relative to the sequence of the wild-type GFP). The substitutions present ce a protein which results in the production of a much brighter the wild-type GFP.

p16-2 promoter was excised from the plasmid "pPD49.78." by digestion HindIII. This 431 bp HindIII-BamHI fragment was inserted between the de #1) and BamHI (nucleotide #31) sites of the GFP-TT gene in the FT," (available from Yishi Jin. at the University of California, Santa igase (Promega). The pGFP-TT plasmid contains the coding regions for into a Tu61 backbone. The DNA sequence of pGFP-TT (SEQ ID NO: 8) re 5. A schematic map of pGFP-TT is shown in Figure 6, in which unique are shown. The pGFP-TT plasmid was digested with HindIII and BamHI, ragment containing the hsp16-2 promoter was inserted, in order to ss-inducible expression vector "pCL25." The DNA sequence of pCL25 and restriction map are shown in Figure 7. Figure 8 provides a schematic in which unique restriction sites are shown.

5

10

20

25

30

EXAMPLE 3

Construction of Transgenic Animals

In this Example, the transgenic parent animals were produced, with one line expressing the β(1-42) minigene (designated as "CL2005") and the other line expressing the hsp/GFP reporter (designated as "CL2070"). For both lines, the transgenes were introduced into *C. elegans* by gonad microinjection as known in the art and described by Mello *et al.* (Mello *et EMBO J.*, 10:3959-3970 [1991]). Marker plasmid pRF4 containing Mello *et al.* (Mello *et EMBO J.*, 10:3959-3970 [1991]). Marker plasmid pRF4 containing the gene (rol-6[su1006]) (SEQ ID NO:9)(pRF4 carries a 4 kb *Eco*RI fragment of *C. elegans* genomic DNA containing the rol-6[su1006] collagen gene in the Bluescribe vector [Stratagene]: *See.* Mello *et al.*, *supra*; and Kramer *et al.*, Mol. Cell. Biol., 10:2081-2090 [Stratagene]: *See.* Mello *et al.*, *supra*; and Roller transgenic progeny were recovered, into morphologically wild-type animals, and Roller transgenic progeny were recovered. The (rol-6[su1006]) gene (SEQ ID NO:9), contained within the pRF4 plasmid is a mutated The (rol-6[su1006]) gene (SEQ ID NO:9), contained within the pRF4 plasmid is a mutated "Roller" phenotype. Figures 9 and 10 show the DNA and amino acid sequences of the rol-6 gene used in this Example, respectively.

Transmitting lines were established and maintained by selection for the Roller marker phenotype. Transgenic animals produced in this manner maintain the injected marker phenotype. Transgenic animals produced in this manner maintain the injected marker phenotype. Transgenic animals multicopy array of variable mitotic and meiotic stability. DNA as an extrachromosomal, multicopy array of variable mitotic and meiotic stability. Strains containing chromosomally integrated transgenes were recovered by irradiation of lines containing extrachromosomal transgenic arrays with 7000 rad (1 rad = 0.01 Gy) of lines containing extrachromosomal transgenic arrays with 7000 rad (1 rad = 0.01 Gy) of lines were then screen for 100% transmittance of the marker transgene. The transgenes in both the CL2005 and CL2070 lines were chromosomally integrated, and were 100% stable.

Transgenic (i.e., as indicated by the Roller phenotype) animals produce both transgenic (i.e., Roller) and non-transgenic (i.e., non-Roller) progeny. These non-transgenic progeny were found to serve as good internal controls for phenotypic and immunohistochemical comparisons. The expression of GFP in the dual transgenic animals immunohistochemical comparisons. The expression of animals from 16 to 25.5°C, at all can be detected in less than 24 hours after the upshift of animals from 16 to 25.5°C, at all

PCT/US97/23819 WO 98/28971

stages of development, from late embryonic to adult. Expression of GFP resulted in the production of green fluorescence in muscle cells; intense tissue-specific expression of GFP was observed. The fluorescence can be observed using compound or dissecting epifluoroescence microscopy (*i.e.*, with standard fluorescein excitation and emission filters). It is also contemplated that the fluorescence is observable by use of fluorimeters and cell sorters.

EXAMPLE 4

Immunohistochemistry of Transgenic Animals

10

15

5

In this Example, immunohistochemistry was used to confirm the transmittance of the chimeric constructs in large populations of putative integrated lines.

As described by Link (Link, 1995, supra), whole mount specimens were prepared by fixing animals in 4% paraformaldehyde and permeabilizing them with 2-mercaptoethanol collagenase as known in the art and described by Link et al. (Link et al., Genetics 131:867-881 [1992]). Three antibodies were used (polyclonal rabbit anti-β peptide antibody, Boehringer Mannheim; mouse monoclonal antibody Ab 4G8, available from Senetek; and monoclonal 4.1, a gift from B. Cordell at Scios Nova). The monoclonal 4.1 recognizes residues 8-15 of β-peptide.

20

25

30

As was observed by Link (Link, 1995, *supra*) animals transgenic for the *unc*- $54/\beta(1-42)$ minigene construct contained muscle-specific deposits of anti- β peptide immunoreactivity. Non anti- β -immunoreactivity was observed in control animals (*i.e.*, non-transgenic, wild-type *C. elegans* tested concurrently). Although *C. elegans* is reported to contain a homolog of the β -amyloid precursor protein gene (*See*, Daigle and Li, Proc. Natl. Acad. Sci., 90:2045-2049 [1993]), this sequence does not contain an apparent β -peptide domain. Thus, it would not be expected, nor did it show, cross-reactivity with the anti- β antibody used in these experiments. For dually transgenic animals, immunoreactive deposits accumulated in the body wall of the animals.

Next, to determine whether the immunoreactive deposits observed in the $unc-54/\beta$ -(1-42) strains displayed the tinctural properties of classic insoluble β -amyloid, transgenic

.... - 001

PCT/US97/23819 WO 98/28971

strains were fixed and stained with thioflavin S, a fluorescent amyloid-specific dye, as known in the art, and described by Guntern and Bouras (R. Guntern and C. Bouras, Experientia 48:-10 [1992]). Thioflavin S-reactive deposits were found in all strains containing the $unc-54/\beta-(1-42)$ minigene constructs, but not in control wild-type animals.

5

Whole mounts of fixed dual transgenic animals stained with coumarin-phallacidin (a muscle-specific probe) and anti- β peptide antibody. Muscle-specific β peptide deposits were observed, and showed a qualitative correlation with the level of GFP expression in these muscle cells (i.e., cells that contained more β -peptide deposits were more green than cells with less β -peptide deposits).

10

In addition, Congo Red and Chrysamine G, two dyes known to interact with β -amyloid and have been reported to interfere with its aggregation are tested on the dually transgenic animals. Anti-oxidants (e.g., vitamin E and ascorbate) are also tested. In these experiments, interference with aggregation of β -peptide, or interaction with β -amyloid are observed.

15

20

All publications and patents mentioned in the above specification are herein incorporated by reference. Various modifications and variations of the described method and system of the invention will be apparent to those skilled in the art without departing from the scope and spirit of the invention. Although the invention has been described in connection with specific preferred embodiments, it should be understood that the invention as claimed should not be unduly limited to such specific embodiments. Indeed, various modifications of the described modes for carrying out the invention which are obvious to those skilled in molecular biology or related fields are intended to be within the scope of the following claims.

- ALLET /DIII E 261

CLAIMS

What is claimed is:

5

- 1. A transparent animal comprising:
 - (a) a toxic transgene:
- (b) an inducible promotor operably linked to a reporter gene, wherein the toxic transgene induces the promotor and wherein the expression of the reporter gene is detectable *in vivo*.

10

- 2. The transparent animal of claim 1, wherein the inducible promoter operably linked to the reporter gene is stably integrated into the genome of the animal.
- 3. The transparent animal of claim 2, wherein the inducible promotor is induced by disruption of cellular function.
 - 4. The transparent animal of claim 2, wherein the inducible promoter is a heat shock promoter.
- 20 5. The transparent animal of claim 2, wherein the toxic transgene is β -peptide.
 - 6. The transparent animal of claim 2, wherein said transparent animal is selected from the class Nematoda.
- 7. The transparent animal of Claim 6, wherein said transparent animal is Caenorhabditis elegans.

WO 98/28971 PCT/US97/23819

8. The transparent animal of Claim 2, wherein said genome comprises SEQ ID NO:5 and SEQ ID NO:8.

- The transparent animal of Claim 2, wherein said reporter gene is green fluorescence protein (GFP).
 - The transparent animal of claim 1, further comprising a heterologous gene marker.
- 10 A method for producing a dually transgenic non-human animal comprising:
 - a) providing:

20

25

- i) a first and second non-human animal; and
- ii) a first toxic transgene; and
- iii) a second transgene, comprising a reporter;
- b) introducing said first toxic transgene into the genome of said first non-human animal to produce a first transgenic animal, and introducing said second transgene into the genome of said second non-human animal to produce a second transgenic animal; and
 - c) mating said first transgenic animal with said second transgenic animal to produce a dually transgenic animal, wherein said toxic transgene and said reporter are expressed.
 - 12. A dually transgenic non-human animal produced according to the method of claim 11, wherein the toxic transgene is β -peptide.
 - 13. A dually transgenic non-human animal produced according to the method of claim 11, wherein said non-human animal is transparent.

WO 98/28971 PCT/US97/23819

5

10

15

20

25

The dually transgenic non-human animal of claim 13, wherein said non-14. human animal is a nematode. The dually transgenic non-human animal of claim 14, wherein said 15 nematode is Caenorhabditis elegans The method of claim 11, wherein said reporter is green fluorescent protein. 16. The dually transgenic non-human animal produced according to the method 17. of claim 11, wherein said first transgene comprises pCL25. The dually transgenic non-human animal produced according to the method 18. of claim 11. wherein said second transgene comprises pCL12. The dually transgenic non-human animal produced according to the method 19. of claim 11, further comprising a heterologous marker gene. A method for testing compounds for toxicity, comprising: 20. a) providing: i) a dually transgenic non-human animal expressing a toxic transgene and reporter gene operably linked to a promoter inducible by the toxic transgene; ii) a composition comprising a test compound in a form suitable for administration such that said compound is bioavailable in the cells of said nonhuman animal; and b) administering said test compound to said non-human animal.

WO 98/28971 PCT/US97/23819

21. The method according to claim 20, wherein the toxic transgene is β -peptide and the reporter gene is GFP.

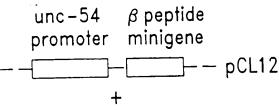
5

- The method of Claim 21, further comprising c) measuring a reduction in the fluorescence of said non-human animal and thereby identifying a compound as therapeutic.
- The method of Claim 20, wherein said compounds inactivate said β-peptide
 expressed by said dually transgenic animal.

÷ : .

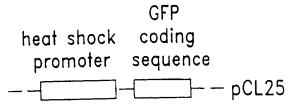


DNA injection mixture 2:



rol-6 mutant collagen gene pRF4

> Microinject DNA into wild type animals



rol-6 mutant collagen gene pRF4

> Microinject DNA into wild type animals



Establish heritable line

CL2005 - muscle specific expression of human β peptide



Establish heritable line

CL2070 - stress-inducible expression of Green Fluorescent Protein



Mate lines together, recover dual transgenic animals by classic genetic techniques

Fig.

Fig. 2

	3/	68	AC.
11	I 160	240	11 11 Hinf 1 16 320 AC 320
	Mbo II Mo II Mbo II Mbo II Mbo II Mbo II Mbo II Hinf Mbo II Hinf Hinf Mbo II Hinf Hinf	B91 I A1WN I FOK I BSP1286 I NIG IV STON I FOK I BGN II GACCATCACTGCCATTGGGATGGCA GACCATTGGGATGCGA GACACATTGGGATGCGA GACACATTGGGATGCGTACCGTAC	Mbo II Mbo II Mbo II Mbo II Hbo II Hbo II THHI

lig. 2A

		4/68	1	2B
243 261 285 291 299 305 316 310 316 310 311 311 311 311 311 311 311 311 311	MDO II TOBI TOBI BS+B I MSO I BSm. L H H H H H H H H H H H H H H H H H H	MSe I Hpg I Hint I Hinc II Hinc III	HE I TTAGGTCGTAAATGI AATCCAGCATTTACA	

VVO 30/2077	5/6	58	1	2C
	640	720	800	Fig.
519 519 520 520 520 520 520 520 520	Sau3a I Mbo I Dpn I Bsty I Bst	Mbo Hinc II RSG I FCOR I DPN PPNACTETTCTGATTCATTTTTCGACCGACCGTACTCTTCGAATTCATTTTTCCCGATCGAT	Scrf I He Ecor II B B	22

6/68		_2D
880	096	Fig.
Scrf I Not I Had I Not I	I Mn! I Mn! I Mn! I Sec I Ban I Fok I Sec I FOK I Sec I Ban I Fok I Sec I FOK I Sec I Ban I Fok I Sec I FOK I Sec I Ban I Fok I Sec I FOK I Sec I Ban I Fok I Sec I FOK I Sec I Ban I Fok I Sec I FOK I Sec I Ban I Fok I Sec I FOK I Sec I Ban I Fok I Sec I FOK I Sec I FOK I SEC I SEC I FOK I SEC	
ACT ACT A TGA T	A L U A G C	l

-- /0111 - 061

	7/68	2 E
I Xmn I StB I StB I Xmn I P-up CAAAGGTATTCGAATGATACTAACATAGAA 1040 CAAAGGTTTCCAAAGCTTACTATGTATTGTATTGTT 1038 1014	RSQ I NIQ IV KPD I BQD I I BQD I I I I I I I I I I I I I I I I I I I	Mbo II Fig.
Hae III Ava II Eae I PpuM I Eae I PpuM I Eae I FC0109 II EC0109 II EC0109 II EC0109 II ECCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC	NSI I HAMATGCATAAGGT TTTACGTATTCCA	1
Ava I Sec I Mn! I ATCAACCTCGGGAT TAGTTGGAGCCCTA 1981 981	2 2 2 2 2 2 2 2 2 2 2 2 3 3 4 4 4 4 5 6 6 6 6 6 6 7 7 7 7 7 7 7 7 7 7 7 7 7	1056 Dde Dde I Hinf I
Stan Stan GAGTAGTGAGCGTAG GAGTAGTGAGCGTAG	Saugana Nia Ava Ppum Ppum Ecool Mnl I Nnl I 1050 1050 1051 1051 1051	EcoR SfaN I

	8/68		2F
1200	1280	1 1360	Fig.
AGATGTGGGTT TCTACACCCAA	Fnu4H NIG III NIG III NIG III NIG III FILL HgiA I Hge II HgiA I Gdi II BSP1286 I E Gd I II BSP1286 I E GG I II BSP1286 I E GG I II FILL RGT II RG	NIG III HgiA I HgiA I NSPH I BSP1286 NSPT524 I BSP1286 STIACTCTTCAACATCCCTACATGCTCTTTCTCCCTGTGCTCCTCTTTTTTTT	

	9/6	o8	ı	つる
1520	1600	1680 680	1760	Fir 2C
\ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \	1499 1505 169 Rsa I ICTGTGTACACTTCTTATGTTTTTTTTACTTCTGATAAATTTTTTT 16ACAÇATGTGAAGAATACAAAAAAAAAGAGACTATTTAAAAAA	MBO III Nde I Mae III Nde I Mae III Nde I Mae III Nde I Mae III Nde I Nde I Nde III Nde I Nde IIII Nde IIIII Nde IIII Nde III Nde IIII Nde III Nde IIII Nde IIII Nde IIII Nde IIII Nde IIII Nde IIII Nde III Nde IIII Nde III Nde IIII Nde	MNI I MAG II BSPH I MAG II DB III ANG II U96 I AGT II NIG III CCTCTCATGACGTCAATCATGCTCATCGT	9

	10/68			-2H
1840	1920	2000	2080	Fig.
	Mbo II Squ3A I Squ3A I HgiA I Mbo I Bsp1286 I Dpn I TCACAAGTATTGATGAGCACGATGCAAGAAGATCGGAAGAAGCCC AGTGTTGATATGAGCACGATGCAAGAAGATCGGAAGAACCC AGTGTTCATAACTACTCGTGCTACGTTCTTTCTAGCTTTCTTCCAAACCC AGTGTTCATAACTACTCGTGCTACGTTCTTTCTAGCTTTCTTCCAAACCCC AGTGTTCATAACTACTCGTGCTACGTTCTTTCTAGCTTTCTTCCAAACCCC AGTGTTCATAACTAGTGCTACGTTCTTTCTAGCTTTTTTTT	B68 1874 1885 1890 M M IGGAGTAGTGTATGGGGTTTTTGCCT ACCTCATCACAGATACCGGA	MSe I 921 921 Fro I Mse I Tithill I Dra I Mse I ACCAAACATATTGAAATTTAAACATTTTCTAAATTTCTTTTAAATTTGCAA ACCAAACATTGAAATTTAAACATTTTTCTAAATATAAAAAATTTAAACGTT ACCAAACATATGAAATTTAAAAAAAAAA	HinP I Bstu I

-- (0111 F 06)

		11/68		21
2160	2240	2320	2400	Fig.
Ecor I Hha I Mae III THAITCIGCATTATTGTGTTTTCCGGCTATATTAATAGGTATTTGTGTTTGTGTTTTCCGGCTATATTAATAGGTATTTGTGTTTGTGTTTTCCGGCTATATTAATAACAAACA	Hinf Hinf Hinf TATGATIC ATACTAAG	HOB III HOB III HOB I EQUI EQUI BQI I BQI	Mae III Aha I Aha I HinP I Hha I Hha I Hallatititisgaaatcatcata AAAIGICGAAAATGAAATCAAGAATAATTCGGCCTTTTTTTTTT	2334

	12/68	,	23
0	7524 7524 256 8 8 59	1	Fig.
6A6616	CACACACACACACACACACACACACACACACACACACA	p8 11	
0 I CGCTA 6CGAT 68	Mn I ACCTCT TGGAGA 2550	Hgd I S+U I OP I D I	
H9 GCCAGAGA CGGTCTCT 24	Hph I GGTGAAA CCACTTT 1 2542	#±±-	-
GCAAAA CGTTTT	h I TGATGAC ACTACTG 33		1
TTTAGG AAATCC	HPF TTTCGG AAAGCC/ AAAGCC/ 25:		
	HH 00 TO	SSCANOR - THE	
CTGTAT]	TCGTCT AGCAGA	S T OK I	
GT CA(CA(CA(CA(CA(CA(CA(CA(CA(CA(CA(CA(CA(C	25 128 11 11 11 11 11 11 11 11 11 11 11 11 11	509 2510	
AGTTATA TCAATAT TCAATAT SPI I SPI I I I I H I I I E	1 B B B C C C C C C C C C C C C C C C C	2505 2505 Alu I	- 1
33.7 M S e T T T T A A A A T T T A Z 4 1 S S G d	E G G G G G G G G G G G G G G G G G G G	Мае I	
233 233 233 1111AAA 1111AAA 103 1 0 1 0 1	2 4 A C C A C C C C C C C C C C C C C C C	889 899 249 F I I I	1
AF. DAG-0	CCCCC → BBS + CCCCC → ABS + CCCC → ABS + CCC → ABS +	NC: NC: NC: NC: NC: NC: NC: NC: NC: NC:	-
AACA	CGTT	. ra.	

		13/68		-2K
	2640	2720	2800	Fig.
1	56AACAGACATTCGCCTACGCCCTCGTCTGTTCCCGACAAGGCAACAGACATTCGCCTACGCCCTCGTCTGTTCCTTCGACAAGGCAACAGAACAGAACAGAAACAGAAACAAAAAA	ATCAGAGCAGATTGTACTGAGAG TAGTCTCGTCTAACATGACTCTC TAGTCTCGTCTAACATGACTCTC AGTCTCTAACATGACTCTC AMSW I AFILII	Sfi I Sau96 I Sfi I Sau96 I FNU4H I Ec00109 I TACCCCATCAGGCGCCTTAAGGGCCTCGT ATGGCGTAGTCCGCGGAATTCCCGGAGCA 2750 2759 2768 2750 2757 2760 2767	88+0

VV () 70/2077 (
		14/68		
1		0	0	\sim
880	096	3040 I	312	
7	7	3 A	ς α ασα	8 <u>F</u>
HA	1 GTA CAT	()()	WAT A CTANTON	- ;
AAA	I GA CT 3	GAA	T P P P P P P P P P P P P P P P P P P P	
1 T A A F	Mbo PGGAA TCCTT 295	I CCA GGT	1 8 1 000 000 000	→
1GT ACA	∢⊢•	Hph TCAC(AGTGC)	3 - 0 3 - 0 3 - 0	
 A T T T A A	AAA TTT	H 6CT(CGA(13(CAA GTT	Hin
00 PT	I TTG AAC	A A A	CT I I I I I I I I I I I I I I I I I I I	-
AC C C C C C C C C C C C C C C C C C C	SP ATA TAT 940	⊢ ∢ ∪	MBO MB MBO MBO MBO MBO MBO MBO MBO MBO M	
1 I I GGA/ GGGA/ CCCT · · · · · · · · · · · · · · · · · ·	ATA TATA 2	O T O O	I E	1
1 n n n n n n n n n n n n n n n n n n n	TCA AGT	TTC	9 GA CT	l L
CACCCA 228	GCT	၁၁၁	HH-HA- W	
AT.	AAT TTAT	T T T A A A	de I TTAC AATG	Z d Se
GAA CTT	ATA TAT	4H I GCAT CGTA	000 000 000 000 000	0 1
9999	A C	00 00 00	1 CA 1	1
TIC	AT .	TTG AAC	1 1 0 0 0 0 0	HgiA
CTT GAA	ATA TAT	T T T A A A	giA Sp12 PdL TGCA ACGT 075 075	
CGTA	I ACA/ TGT	CCT	TBA-00-NWM	1
00 P	III I GAG/ CTC	8 TTC AAGG	I TTG	
CAG GTC	TAT TAT	90 1 A T		11
CGT CGT 24 24	ထ ပပ		Sau Mbo II Dpn CCTA CCTA 306	Xmn I
AD T T T T T T T T T T T T T T T T T T T	ပ္ပိမ္တ	T C G C	Mbo GAA CCTT	칟
Dde CTT GAA 282	TAT ATA	A C	7 00 9	1 1
TTT(A T C C	ວງ9) ວວວ	Sfal GAT CTA 305	
166°	AAT TTA	TTT	A A A A T T T	1
TAATT	TCA AGT	ACA	• GTA.	
AA + .	CAT GTA	TCA AGT	AA T	
III I I I I CTA	ATA TAT	TAT	oh I STGA SACT	1 1
NIG NIG SpH CAT (GTA 2802	TAA/ ATT	GA CO	H H P P P P P P P P P P P P P P P P P P	
CG1-B	ပ်ဖွဲ့		ပၒ	

	15/68	3	$^{-}$ 2M
3200	3280	3360	Fig.
CCTTGAGAGITTTCGCCCCGAAGAAGGTTACTAGTCGTGAAAATTTCAAGACGATACACGCGCGCG	T ap = 000	3210 3210 3210 3210 3210 5faN I Fok I NIG III SCATCITACGGAIGGCAIGAGAGAATTAIGCAGIGCTGCATAACCAIGA 5GTAGAAIGCCTACCGTACTGICATICICTTAATACGICACGGGTAIIGGIACT 33326 3337	Sau3A I Mbo 1

	16/68	3	$\frac{1}{2}$
3440 0 0	1 3520	3600	Fig.
MDD DDD DDD 344 M444 M444	Mde I CAACG GTTGC 3518	TAAAGT ATTICA	1
111 141 1428	AATGGCAA TTACCGTT	Mn	
22 22 34 25 34 25 34 25 34 25 34 25 34 25 34 25 34 25 34 34 34 34 34 34 34 34 34 34 34 34 34	161AGC ACATCG	Fok I GGATGG GCTACC 3582	Hph I dsw
1 00	Sfan I CGATGCC GCTAAGG	se I RATAGACI ATTATCTG/ 71	1
N19 16CACAACATG ACGTGTTGTAC 341	10 III 10 ACACCA ACTGTGGT 486	AS AS AS AS AS AS AS AS AS AS AS AS AS A	
GCTTTTT	CGAGCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG	Msp Hpg Scrf Nci I Bcn I Bcn I 1000000000000000000000000000000000000	വധ
I TAACC ATTGG	ACCAAACGA(,TGGTTTGCT	Alu Mae I 1 1 ACTCTAGCT ATGAGATCGA 3554 3556	,
I I A I I COTT CCT T CCT	SCCAT CGGTAT	TACT1	MSP 1 96
Sau96 Ava I Ava I GAGGACC GTCCTGG 1383 3383	3381 I GAATGAA	GGCGAAC	Sau9
Mbo 1 Dpn I Pyu I	23.7 20.1 20.2 20.2 20.2 34.2 34.2 34.2	452 452 MSe TAATT	
TCTGACA AGACTGT	NIG H H - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1	CAAA GTTT	1
CTTACT	A T C C C C C C C C C C C C C C C C C C	Hin Hin Hin September 1982 A A C C C A A C C C C A A C C C C A A C C C C A A C C C C A A C C C C C A A C C C C C A A C C C C C A A C	5

	17/6	8	20
3680 3680 80	3760	3840	Fig.
061CTC CCAGAG 36	GCAACT CGTTGA	TTACTC AATGAG	D N
AGCGTGC TCGCACC	le I inf I AGTCAG TCAGTC 748	CCAAGT GGTTCA	ł
HPG III 11V 11V 11V 11V 16GCCAC 1662 3663 3663 3665	ACGGGG - H GCCCC 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3	I GTCAGA CAGTCT	
N D N D C C C C C C C C C C C C C C C C	CTACACG GATGTGC	MGB II TGGTAACT ACCATIGA 3817	Squ3A Mbo I Dpn I A N I BS + Y I
GATAAAT	AGTTATO •	1 AGCAT TCGTA 7	A I Mb I Mb
ATTGCT	TATCGT ATAGCA	MSB TGATTA SACTAAT 380	Sau3 Mbo Mbo Mbo I Dpn
TGGTTT	Mn l I CCTCCG GGAGGGC 3717	Mn I I I I I I I I I I I I I I I I I I I	Mse Dra I
GGCTGGC CCGACCG 29	GTAAGC CATTCG	N 1 Q B B B B B B B B B B B B B B B B B B B	
111 Hp	III 6 I V CAGATG GTCTAC	D d d d d d d d d d d d d d d d d d d d	3/86
0 1 1 1 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	Hae Nau9 Nau9 Nau9 Nau9 Nau9 Nau9 Nau9 Nau9	. JOE HA 800	æ ₹
Hin Hin Hha Hor Hha Hada AAGACGC AAGACGC 361	Fnu4H Bbv I GCAGCA(CGTCGT(3692	AAATAG TTTATC	
GACCAC CTGGTG 605 605	ATCATT TAGTAA	k I ATGAACG TACTTGC 63	
ACCIA ACCIA ACCIA ACCIA	GCGGT	Fok ATGGA TACCT 376	

WO 98/28971			
	18/68	<u>0</u>	J
3920	4000	1 4080	$\frac{18}{6}$
	Pn I Mbo II Dposty II	Tthill II NSPB II Tthill by I Tthill II CIGCTTGCAACAAAAAACCACCGCTACCAGCGCTGGTTTGTTT	4020

-- /--

		1	19/68			20
 	e III 4160	09	4240	4320		Fig.
	HQB HQB GTTAG CAATC		GTGGC CACCG	6611C CCAAG	1	
7 7 069 4070 4070	ပ္ပို့	nu4H I H I bv I	CTGCCA GACGGT 229	. ⊃Ö		
406 406 4	e I AGTGTA TCACAT	2 Fnu4 I B Bbv	TGCTG ACCGAC 4226 1 4226 1 4226	GGCTGA		
	MQ CCTTCT GGAAGA	∀ ⊢-	TACCAG ATGGTC 16 422	H I I I I I I I I I I I I I I I I I I I		
	TACTGT	X -	TCCTGT AGGACA 42	NSP FNU4 Bbv HinP I Hhd I	4 2 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4	
	ACCAAA	•	TGCTAA ACGATT	GATAAG	, ww	
	a I I I C C C C C C C C C C C C C C C C	. ### _ I	TCGCTC AGCGAG 98	MSP HPG TTACCG	279 279 42 42	
	Hi Hh CAGAGC GTCTCG	- 44 Z-	CATACC GTATGG 41	CGATAG		1
	CTTCAG	•	CGCCTA	I f I TCAAGA) - •	
20 - — 4022	e III Aactgg Ttgacc	00	TAGCAC ATCGTG	I Ple Hin TIGGAC	7 7 7 - 44 - 60	
40	Ma GAAGGT CTTCCA	- 4	ACTCTG TGAGAC	NCCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG	2-4444 2-4222 2-2555 3-7-7-7-7-7-7-7-7-7-7-7-7-7-7-7-7-7-7-7	
4010	T T T C C		TCAAGA AGTTCT	110191	A > 4 • • • • • • • • • • • • • • • • • •	1
4	AACTCT TTGAGA	ر ن نون	ACCACT TGGTGA	A A G T C G	- CAG	
	ACC	2	999 999	GAT	_	

	20/68		2R
4400	4480	4560	I I Fig.
CGCCA GCGGT 392 392 392 560 SGC	D TA O	I STCGTCAGG SAGCAGTCC	N N D I I N N S D H I N N S D 7 S 2 4 A 1 I I I I I I I I I I I I I I I I I I
ATTGAGAAATTAACTCTTT	Mnl I I Alu ACGAGGAGC TGCTCCTCG TGCTCCTCG	SfaN TTTGTGATGC AAACACTACG 4548	Hae III
CAGCGTGAGC GTCGCACTCG	HinP Hha Hha TCCTCTCGCG 1459 4459	Taq I Hga I GAGCGTCGATT CTCGCAGCTAA 4534	
I AGATACCTA TCTATGGAT	GGTCGGAAC CCAGCCTTG	Mn! I CCTCTGACTTG GGAGACTGAAC 4522	HDE SCRF I ECOR I BS +N I
CACCGAACTGGTGGTGGTGGTGGTGGTGGTTGACTGAC	I Fnu4H I II HAGGGGGAGATTGGCGGGGGGGGGGGGGGGGGGGGGGGGG	TTTCGCCA AAAGCGGT	Hde III nu4H I +U I
AACGACCTAC	MSP HPG CCATAGGCC 4430 4430	CCTGTCGGG	88
u I CTTGGAGCG GAACCTCGC 34	AAGGCGGACA TTCCGCCTGT	TCTTTATAGT AGAAATATCA	
6 I A I A I A I A I A I A I A I A I A I	AAGGGAGA TTCCCTCT	SCTF I ECOR II BS+N I CCCCTGGTA CCCCTGGTA 4488	48 - I
Hgia I Bsp128 Apgl I CACGC CACGTG 4321 4321	TTCCCG AAGGGC	GGAAA(CCTTT	

	21/	68	
4640	4720	4800	Fig.
GGGGCGGAGCCTATGGAAAAACGCCGGCCTTTTTACGGTTCCTGGCCTTTTGCTGGCAAAACGAGGAGGGCGGAAAACGAGGGGCGGAAAACGAGGGGCGGAAAACGAGGGGCGGAAAACGAGGGGGTGTACAAACGGGGCCGGAAAACGGGGCCGGAAAACGGGGCCGGAAAACGGGGCCGGAAACGGGGGTGTACAACCGGGAAACGGGGCCGGAAACGGGGGTGTACAACCGGGAAACGGGGCCGGAAACGGGGGTGTACAACCGGGAAACGGGGGTGTACAAACGGGGCCGCGCGCG	Fnu4H I Hinf I Hinf I AGAAAGGACGCATTACCGCTTTGAGTGACTGATACCGCTCGCCGCAGCCGA AGAAAGGACGCATAAGACGTATTACCGCCTTTGAGTGACTGATACCGCTCGCCGCAGCCGA AGAAAGGACGAATAGGCGACTATTGGCATAATGGCGGAAACTCACTC	T-00-4 VV	Mse I Alu I Ase I Nsp8 II

	22/68	2T
4880	4960	Fig.
Hinf I Bby I Ase I Ase I Alu I Ase I Ase I Alu I Ase I	8813 813 813 813 S S S S S S S S S S S S S S S S S S S	4973

23/68 4813 Pvu II 3522 Fsp I 3375 Pvu I 2488 BamH I 3264 Sca I 2940 Ssp I 2762 Afl II 2509 Apa I 2504 Spl I 2494 Spe I 1252 Sac I 1119 Kpn I 1119 Asp718 1116 BspM I 476 Hpa I 1071 Nsi I 415 Pst I 1061 Nhe I 375 Bsm I 1056 Sty I bv II 833 Sma I 833 Xma I 804 Mlu I 16 HinD III 24 Bbv II

Fig. 3

Fig. 4

gly val val ile ala GTC ATA GCG GCA TTT GCA GAA a l a ACC GAT thr asp phe gin lys leu val phe 51 gly 110 ATC TIT CTG GCA CCA GCA GGT GTG GGC GGT GTT <u>a l a</u> CAT CAT CAA AAA TTG GTG pro gly leu met val gly a l a ile phe leu c 1161 / 31 GGT TCA AAC AAA GGT GCA ATC ATT GGA CTC ATG gly ser asn lys gly ala ile ile gly leu met his his g 1221 / glu val 111 ' met his lys val leu leu ala leu phe phe 1131 / 21 GAA TTC CGA CAT GAC TCA GGA TAT GAA GTT AAG GIT TIG CTG GCA CTG TTC tyr arg his asp ser gly 41 I gly ser asn lys g 61 glu phe 1191 / GAT GTG asp val 1071 /

Fig. 5A
Fig. 5B
Fig. 5C
Fig. 5D
Fig. 5E
Fig. 5F
Fig. 5G
Fig. 5H
Fig. 5I
Fig. 5J
Fig. 5K
Fig. 5L
Fig. 5M
Fig. 5N
Fig. 50
Fig. 5P
Fig. 5Q
Fig. 5R

Fig. 5

26/68 80 atgittegaatgataet Itacaaagettaetaiga 68 69 Tag Bs+B Sau96 NIa IV Hae Hae Eae Bal NOC. I SECOND SE Mbo Dpn N-d Ba+Y BamH 25 26 Ple Hinf 22 22 Sa თთთ Sph I NSph I NSpH I NSp7524

	27/68	0 g. 5B
34 33 34 1 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3	IV KPD I II BDD I 109 I ASPZI 100 I MNI I H 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	110 110 110 HPh I MS GATGGTGATGTT TCTACCACTACAA 191 19

28,	/68	2C
320	400 1 480	Fig.
Habiti H	Squ3A I Mbo I Mbo I Ppn I CCACAAGTTACGAAAGGTTCTATGGAACGGCATGACTTT CCACAAGTTACGAAAGGTTCTATGGATCATTTGCCGTACTGAAA 359 359 359 359 359 Af III ACAGGAAGAACTATATTTTCAAGGTGACGGGAACTACAAGACACGTGC	

	29,	/68		(7
560	→ Thr 167	720		800	ب ا
MSe I Hinf I MSe I Hinf I MSe I Pra I APA I GATACCCTTGTTAATAGAAGATTGAATTTGAATTTAAAGAAGATGGAAACATTCTTGGACACAATTAACTTAATAGAAATTTCTTGTAAGAACCTGTGTTTAACCTTA AGAACCTTTAAGAACCTTTAAGAACCTTTAAGAACTTAACCTTAAGAACTTAACCTTAAGAACTTAAGAACTTAAGAACTTAAGAACTTAAGAACTTAAGAACTTAAGAACTTAAGAACTTAAGAACTTAAACCTTAAAGAATTAACCTTAAAGAATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAAGAAATTAAAGAAATTAAAGAAAATTAAAAAA	ACTATAACTCACACAATGTACATGTACGTGCAGACAAAGAATGGAATCAAAGTTAACTTCAAAAGTATGAAGTTTTGAAAGTTTTTTTT	MGE I CAACTAGCAGACCATTATCAACAAATACTCCAATTGGG GTTGATCGTCTGGTAATAGTTGTTTTATGAGGTTAACCC	Squ3A I Mbo I Dpn I A+w I Bs+Y I	AGACAACCATTACCTGTCCACACAATCTGCCCTTTCGAAAGATCCCAACGAAAAGGAGGGGGAGGACCACGTGGTCTTGTTGAGGTTCTCTTGGAGAAGGAGGAAGGA	

	30/68		-5E
	880	096	Fig.
754 787 787 761 761 761 761 761 761	Note 1	ACCAACTIGICTGGTGTCAAAAAI	NIQ III

	31	/68		5F
1040	1120 1120 20	1200	1280	Fig.
524 I BSP1286 I BSP1286 I GCTCTTTTTTGTTATTATCCGAGAAAGAGGGACACGAGGGTGGGGGGATAAAAACAATAATAGGGAGAAAAAAAA	Hph I Mae III Xma I Xma II Xma II Xma II Mse I Mnl I Xma II Xma II II Xma II Xm	MSE'I WSE'I HGIA I HSP1286 I ASE I TAATTCGTAATAAAAAATCTACACAAATGTGTTACAAATGTAATA TAATTCGTAATAAAAAAAAAA	O I ACACTICITATGITITITITACTICIGATAAITITITITGAAACAICATAGAAAAACO TGIGAAGAATACAAAAAAAAGACTATTTAAAAAAAAACTTIGTAGTATCTTTTTIG	MGE III MGE II AND III AND IIII AND III AND II

W O 98/289 / L				
		32/68		IJ
1		₩	_	5
09	1440	286 520	009	
13.	•	P 1 20 20 20	-	6d: Eag Eae
A → → S	⊢ ∀ • :	AC - BS - B	Co A ·	8 M M
TATC TAG 1	GCT CGA/	GA	A A A A T T T	
CAA	C T G G A C	A T T T A A	P C C	1
61 09 50	AAG	AGT	A A T T T T A A T	
1-00 1-10 1-10 1-10 1-10 1-10 1-10 1-10	∓A •	STA.	A T A T T A T	1
00 A C A A C A A C A A C A A C A A C A A C	TAA ATT 24 425	T C A A G T	A C	
A A A A	- AH- L - AH- L - AH- 41	AAA TTT	AGI	i I
CCT GGGA 134 39	GA	CTA GAT	TAGA	
1 3 3 · CCC 1 3	T A T A T A	A C	GO.	(
CT	GTT	TCT AGA	h A C 7 3	
CGT GCA 332	AA TT	1 T T A A A	HP HP 15 15 15 15 15 15 15 1	1
- 450 	TGA ACT	CG1 GCA7	SGAA CCTT	1
TC6	A A G T T C	00 00 00	1 001 001 001 001	
132 132	AGT.	I GAC TGC	dе 1СА AGT 563	1
A A T	CAA GTT	—H നറ ∞	0H-00-6-	
TTT AAA.	111 AAA	MD T 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	Mn I GAG CTC 155	i
T T A A A	ATT	GTT	1 T T T A A A A A	
CAA	GA CT	GA CT	၅၁ ၅၁၁	
90 99 99	T T G A A C	II CAA GTT	TTT.	
SO TO	TTT AAA.	11 161 ACA	ဗ္ ဗ ပ	\
TAT	T A	1 1 A A A A A A A A A A A A A A A A A A	II GAA CTT	5
CAA	GAG CTC	T + 1 T A T T G	bo I TT	154
OC →	T C	CTA GAT	3A I I I I I I I I I I I I I I I I I I I	
7 T T T A A A A A A A A A A A A A A A A	111 AAA	ပ္ပံ ပို့	Sau Dpn CTAT 153	153
29	AAG.	T T C A A G	A A T	1
AAAA2	A T T T	GGT	AGA	
- TAT - 7	TGA ACT	ဗ္ဗာ	N I GCA CGT	1
A T T A 2 2 8	A GC	AACC	fai AT TA 52	
TAT TAGO	CA	TTT	ν-00 Δ Ε	1
CTT	GCT CGA	ပ္ပံ	A C C C	
1 38	H	⊬∢	~ [-	

_ ._.. _ ...

W 5 70/207	33/68	Н9
I Xmn I Spe	Scrf Scrf Scrf Scrf Scrf Scrf Scrf Scrf Script	Hga I Bstu I Fig. 5
GGAGTAGTGTCTATGGGGTTTTTGCCTTAA CCTCATCACAGATACCCCAAAAACGGAATT	Saugé I NI a IV E COO109 I B Sp 1286 I B Sp 1286 I B Sp 1286 I B Sp 1286 I HinP I He III Hae III BS 10 I Hae III Hae III Hae III BS 10 I Hae II Hae III Hae III BS 10 I Hae III Hae III Hae III BS 10 I Hae III Hae III Hae III Hae III BS 10 I Hae III Hae II	ν (Φ

	34/68		51
1840	1920	Dde I IC 2000 1G 2000 1G 2000	Fig.
ACACTIGICAAGCGGATGCCGGAGCAAGCCCGTCAGGGGGGTGTTGGCGGGTGTCGGGGCTG ACACTIGICTGTAAGCGGATGCCGGAGCAAGCCCGTCAGGGGGGTGTTGGCGGGTGTCGGGGCTG ACACTIGICTGTAAGCGGATGCCGGGAGCAAGCCCGTCAGGCGGGTGTTGGCGGGTGTTGGCGGGTGTTGGCGGGTGTTGGCGGGTGTTGGCGGGTGTTGGGGGG	ω <u>σ</u> Ο	MSB I BSPH I BSQU96 I WIG III WIG III DECOOTO III WIG III DE LE COOTO III DE LE COOTO III DE LE COOTO IIII DE LE COOTO IIIII DE LE COOTO IIIII DE LE COOTO IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII	1942

	35/6	8		5
2080	2160 I	2240	23	Fig
	NIG III BSPH I BSPH I CTCATGAGAGAGAGAGTTGTTGAAAAGGAAGAGGAGGTATGAGTTGAGTTGTGAGTTGT BSPH I CTCATGAGAGATAACCTGATAATGCTTCTCATACTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG	Hph I Hph I 2087 2087 2088 Fnu4H I Fnu4H I 2000 Hph I Hph I 2088 CAGCCCTTTTTTGCCTTCCTGTTTTTTGCTCAGAAACGCTGGTGAAGTAAAACGATGCTACGAGTGGGTCTTTGCTATTGTACGAGTGGGTCTTTGCTATTGTACGAAAACGAAAAAAAA	DPN I BSIN I BSIT I BSTT I GCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC	2242 2245 2245 2245 2245 2245

3	36/68		5 K
2400	2480	I 2560	Fig.
Scrf Nci	I Fok I TACGGATGG	u4H I v I TGCCATAACCATGAGTGATAACAC ACGGTATTGGTACTCACTATTGTG	
Mae II Hgia I Maspila I Maga I Maga I Maga II Bspila I Bspila I Maga II Bspila I Maga II I Bspila I Maga II Ma	Fnu4H I Pd CAACTCGGTCGCCGCATACACTATTCT GTTGAGCCAGCGGCGTATGTGATAAGA 2411 2411	Fn Bb CATGACAGTAAGAGAATTATGCAGTGC GTACTGTCATTCTCTTAATACGTCACG 1 2481	

	37/68		7S
2640	2720	2800	Fig.
Squ96 I Avg II Mni I Mni I Mni I Mni I Mni II M	09 EQ		

38/68	Z Z
Fnu4H I NIG IV Bbv I NIG IV TGCAGCACTG 2880 AACGTCGTGAC 2880 2872 2880 2872 2880 cGTTTATCTGT 2960	MSe I Dra I TITAGATIGATI 3040 AAATCTAACTAA 3039 3040
BSTU I GTGGTCTCGCGGTATCA CACCCAGAGCGCCATAGT 2860 1 FOK I 1 FOK I 2943	AGIITACICATATATAC
AATCTGGAGCCGTGAGCCTTAGACTCGGCGTGAGCCTTGGAGCCTCGGCCACTCGCACTTGGAGCTTGGCTTTAGACCTCGGCCACTCGCTTTAGACTTAGACGACGGGGAGTTAGACGACGGGGGGGG	66666
CTGGTTTATTGCTGATA GACCAAATAACGACTAT GACCGTATCGTAGTT CCTCCCGTATCGTAGTT GGAGGGCATAGCATCAA	M
MSP I HPG II HPG	2882 2882 0u3A I NIO bo I Ban Pn I Ban ATCGCTGAGATAGG TAGCGACTCTATCC 1961 2961 29

	39/68	5N
IDE II ICGTGAG 3120 IGCACTC 3114	I Fnu4H I Bbv I TAATCTG 3200 ATTAGAC 3200 3200	1111 1280 1g.
MSB WSB CCAAAATCCCTTAA GGTTTTAGGGAATT 3111	I BstU HinP Hhd I TTTTTCTGCGCG AAAAAAGACGCGC 3189 3189 3190	MSP I MSP I HPG II AIU I HPG II AIU I HGGACCTACCAACTCTTTTTCCGAAGG 3
NIG I BSPH I I CATATCTCATGA ACTATTAGAGTACT 3095 3095	3A I Sau3A I Bbo II Mbo II Dpn I I A W II I CTTCTTGAGATCC GAAGAACTCTAGG	MSP I HPG II Alu
A [w I B S I Y I I I I I I I I I I I I I I I I I	Sau	B II
AAAAGGATCTAGGT AAAAGGATCTAGGT TTTTCCTAGATCCAG TTTTCCTAGATCCAG 3065 3065 3066 3066 3066 3066 3066 3066	I CAGACCCCGTAGA GTCTGGGGCATCT	NspB LCCACCGCTACCAGC
MSe Drd MSe I H I I I I I I I I I I I I I I I I I I	Hga I Dde I GCTTCCACTGAGCGT(GCAAGGTGACTCGCA(3132	THHIII II CTGCTTGCAAAAAAACCACCGCTACCAGCGGTGGTT
TAAAA(TTTT AAAAA	1+4

	40/68	50
3360	I I 3440	3520 Fig.
GCCACCACTICAAGATICTE 328 111 6 III 6 CCACCACTICAAGAACTCT CCGTGGTGGTGTTCTTGAGA	Scrf I	Hgia I Hinh I Hinh I Hinh I Apal I Apal I Apal I Apal I Apal I Apal I Hinh I HIN I HINH I HIN
•		

	1 4080	13/68 2	J D
7	MSe I Sec I	. ∀ > ⊢∀	\
3942	HinP I Hha I Hogentreceactesaaagegegegegegegegegegegegegegegegegeg	MSP II HPG III CGAAATGIGAATACGAAGGCCGAGCATACAACACACCTTAAG 4099 4099 WID III CCATGATTACGCC 4173 GGTACTAATGCGC 4173	

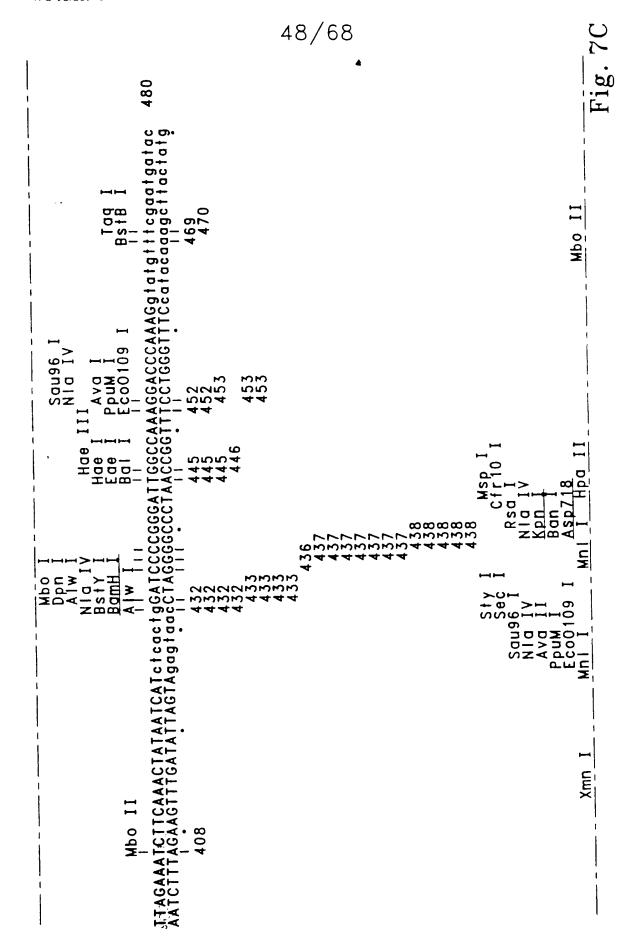
44/68 HinD III 7 Sph I 13 Pst I 15 BspM I 19 Sal I 25 Xba I 31 BamH I 36 Ava I 36 Sma I 36 Xma I 118 Asp718 118 Kpn I 299 Nco I ,326 581 Xca I Ser 65 → Thr 620 Hpa I 633 Ile167→Thr 850 Bsm I 859 EcoR I presumed gGFP-TT 4173 base pairs unique sites Fsp I 2702 Pvu I 2555 Spe I 1674 Sca I 2444 1680 Eag I Spl I 1684 Ssp I 2120 1689 Apa I San II 1689 Fig. 6 1942 Afl II

Fig. 7A
Fig. 7B
Fig. 7C
Fig. 7D
Fig. 7E
Fig. 7F
Fig. 7G
Fig. 7H
Fig. 7I
Fig. 7J
Fig. 7K
Fig. 7L
Fig. 7M
Fig. 7N
Fig. 70
Fig. 7P
Fig. 7Q
Fig. 7R
Fig. 7S
Fig. 7T

Fig. 7

	46/6	8			77
8 I 8 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	—	160		240	Fig.
Ta NIG III BS† 	Fnu4H Pst I Bsm I Bbv I	ATATACCCGCATTCT TATATGGCGTAAGA 149 15	Hgd I	AAAGAGAGGAGAGGAAAA TTTCTCTGCGTCTGCCTTTT 226	
Hinf I GCATTTGAATCAGAATATGGA CGTAAACTTAGTCTTATACCT 46	Bsp1286 I Ban II Dde I Mni I		4	GAAAATACTC CTTTTATGAG) l
Pie I Squ3A I	Wde	ပ်ပ	au96 I UM I Sfa 00109 I Mae I Va II Ns	AGAAGGTCCTAGA TCTTCCAGGATCT	0 0
NIG III Sph I NSPH I NSPH I NSPT524 I NSPT524 I HIND III PST TICGAACGTACGGAC TICGAACGTACGGAC	4	CGAAGTTTTTTAGA GCTTCAAAAAATCTV	∑ £	CTTACAGAATGTTC GAATGTCTTACAAG	

	4	7/68		_7B
I 320		400		Fig.
184 190 209 210 T Alu Fnu4H Bby I Hinf I Hinf I I Hill II Hinf I I I I I I HINF I I I I I I I I I I I I I I I I I I	71G11CG1CG 311 311 314 314 316	MDB I Xbd I 	Scrf I Nci I Msp I Hpd II Ecn I Smg I Sec I Scrf I	Ava I Sec I Squ3A I



	49/68		_7D
aacataacatataaaattttcagGAGCCCTTGGAGGGTACCGGTAGAAAAAAAAAAAAAAAA	511 511 Hph I Ms GATGGTGATGTT TCTACCAA 592 59	o нн oo- o	703 703 704 Fig.

	50/	⁷ 68		7E
800	h I 880 0	4 096	1040	Fig.
CATGC GATGC GTACGG	Hph TTGAAG AACTTC 880	TTGGAA FACCTT	hr167 Mae I CIAGACA GATCTGT 1034	
GAGTGC	TCAAGT AGTTCA	ACACAAA TGTGTTT	TCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	
TTTCAA	I I GCTGAAG CGACTTC	TCTTGG/ AGAACC1	HSB I FIG II FITAACT TTAACT AATTGA 021 021	
ATGACTI TACTGAA75	Mde 11 11 	AAACAT	nf I H ATCAAAG TAGTTIC	
A A C G G C C C C C C C C C C C C C C C	TACAAG	GA CT	H : H : A A T G G A T T A C C T	1
DAHH AH	763 CGGGAAC GCCCTTG	489 I Frd I I TTAAAG TTAAAG AATTTC 22	AAAAG	1
Sau3 Mbo Mbo Dpn CCAGATC GGTCTAG GGTCTAG 760 760	AGATGA TCTACT	D D D D D D D D D D D D D D D D D D D	I T+h11 AGACAAAC TCTGTTTG 999	
AGATAC TCTATG	TTTCAA	I AAAGGTA TTTCCAT	NIG III CATGCC/ GTACCG1	1
TTTTCA	CTATATT GATATA	I Mse q I GAGTTA/ CTCAATI 2 907	d I c I ATACAT TATGTA 2	
TCAATGC AGTTACG	AAAGAA TTTCTT	Hinf - Ta AGAATC TCTTAG 1- 1- 90 899		
TGGTGT	sa I TACAGG ATGTCC	Mse I 1 1 1 1 1 893	ACTCACA TGAGTGI	
Thr65 CACIITA	R GGTTATG CCAATAC 8	TACCCT1 ATGGGA/	ACTATA IGATAT	1
ACTTT TGAAA	CGAAC	CACT	TACA	

WO 98/28971		PCT/US97/23819
	51/68	7.
1120	1200	1 1280 Fig.
MDO II MORE I HORE TECACOTICA ACTION CAACAAATACTCCAATTGGCGATGGCCTGTCCTTTTAC GTTGTAACTCCAATTGGCGCTACCGGGACAGGAAATGGTTGTAACTCCAATTGGCGCTACCGGGACAGGAAATGGTTGTTATGAACTCCTACCGGGACAGGAAATGGTTGTTTATGAGGTTAACCGCCTACCGGGACAGGAAATGGTTGTTATGAACTCCTACCGGGACAGGAAATGGTTGTTTTATGAACTCCTTCTACCTTCGCAAGTTGATCGTTGTTTTATGAGGTTAACCGCCTACCGGGACAGGAAATGGTTGTTTTATGAACTCTTCTACCTTCGCAAGTTGATAGTTGTTTTATGAACTCGCTACCGGGACAGGAAAATGGTTGTTTTATGAACTTCTAACCGCTACCGGGACAGGAAAATGGTTGTTTTATGAACTTCTAACCGCCTACCGGGACAGGAAAATGGTTGTTTATGAACTTCTAACCGCTACCGGGACAGGAAAATGGTTGTTTAACCGCCTACCGGGACAGGAAAATGGTTGTTAACCGCTAACTACCTTTTAACCTTAACCGCTAACTAA	Sau3A I Mbo I Dpn I Alw I Bstw	Hinp I Nob III Nob

52/68	52
1360	$\begin{array}{c} 1 \\ 1520 \\ \end{array}$
NSPB II	AATTTCTTTGTTTTTAGCTTCTTTTAAGTCACCTCTAACAATGGAATTGTGTGTG

1 1600	AT 1680	NIG III	TT 1840	3AT 1920
FA	TA			Fig.
ASB I AST CONTRACTOR AND	RSO I GTICTGIGIACACTICITATGITITITIACITCIGATAAATTITITITIGAAACATCATAGAAAAAAAAACGACACACAAAA CAAGACAÇATGIGAAGAATACAAAAAAAAGAGCTATTTAAAAAAAAAA	Hae Mae III Squys Mae III Squys Mae III Squys Mae III Squys Mae III Harring Mae III Squys Mae III Harring Mae II Harr	MSB I MSB I ASB I ASB I ACGAGTAGCACTTTTTCAAAAACCTCATAAAAACCTTAAAAAGGTTAGTTCACTTTCAATAATTAAT	Whi I Tag I TOCITITE GEOGITIC CCCTATIGITA GAGITIC GAGGACGGCGTTITIC TECTE GEORGE CAAGTATTGAT

		54/68	71
	2000	2080	2160 2160 Fig.
ACGAAAACCCCCAAAGGGGATAACAACAGTTCTCAAAGCTCCTGCCGCA 1865 1865 1880 Mbo II Sau3A I Mbo II Mbo II Mbo II Mbo II Mbo II	GCACGATGCAAGAAGATGGGAAGGTTTGGGTTTGAGCTCAGTGGAAGGTGAGTAGAAGTTGATATTT GCACGATGCAAGATCGGAAGGTTTCTAGGTTTGGAACTCCGAGTCACCTTCCACTCATCTTCAACTTAAA CGTGCTACGTTCTTTCTAGCCTTCTTCCAAACCCAAACTCCGAGTCACTTCCACTCATTAAAAAAAA	AG 76 76 76	Saug6 I Rsa I Squ 96 I Squ 96 I Squ 96 I Squ 96 I Spi 1 Sp

	55/	68	7.3
	2240	2320	Fig.
2140 2143 2143 2145 2145	Hin Bs Hin Hha Hha 1-1 1-1 1-1 1-1 1-1 1-1 1-1 1-1 1-1 1-	75 0	
	2086 2090 u I Fok cttgtctgtaagcggaT	Sfan I	MS9 I Afl II

	56/68		AK
2400	2480	2560 I	2640
BSPH I HATTITE AT A SETT TATA	CCTATTIGITTATTTTCTAAATACATTCAAATATG 24 GGATAAACAAATAAAGATTTATGTAAGTTTATAC 1	TIGAAAAGGAAGAGTAIGAGIATICAACATTICCG 25 AACTTITICCTTCTCATACTCATAGTTGTAAAGGC 2534 Hph I Hph I Sfan I	AACGAGTGGTCTTTGCGACCTTTCATTTCTAC 2609 2624 2637
Sfill Sauge I Sfill Sauge I Fnu4H I Eco0109 I Fnu4H I Eco0109 I Hqe III Hqe III Hqe III Hqe III CTTTTATGGCGTAGTCGGGGCCTTAAGGCCTCGTGATACGC CSTTTATACCGCATCAGGCGCCTTAAGGCCTCGTGATACGCC TCTTTTATGGCGTAGTCCGCGGAATTCCCGGAGCACTATGCGG 2331 2334 2347 2341 2334	II III TCAGGTGGCACTTTTCG AGTCCACCGTGAAAAGC	GAGACAATAACCCTGATAAATGCTTCAATAA CTCTGTTATTGGGACTATTTACGAAGTTATT 25	GGCA CCGT 83

	57/68	7.Z
2720	2800	2880 Fig.
Squ3A Squ3A Squ3A Mbo Mbo	A A A A	RSG I MGE III ŞfGN I FOK I KOLITGGTTGAGTACTCACCAGTCACAGAAAGCATCTTACGGATG
Sau3A I HgiA I Taq Mbo I Bsp1286 I Mae III o III ApaL I Mae III cTAGTCAGCCACGTGCTCACCCAATGTAGC ICTAGTCAACCCACGTGCTCACCCAATGTAGC IS SEA SES SES SES SES SES SES SES SES SE	Mae II Hgia I Mse I II II BSP1286 I BSP1286 I ACCTTTTCCAATGATGAGGTTACAAGT TGCAAAAGGTTACTACTCGTGAAATTTCA 2741 2749	ACTCGGTCGCCGCATACACTATTCTCAGAATG TGAGCCAGGGGCGTATGTGATAAGAGTCTTAC
NAS C C A A A A A A A A A A A A A A A A A	22 - 4C	CTC

			58/68		Z.
1		2960	3040	3120	F. 0
	Sau Sau	ACT ACT	Squ3A I Avg II Avg III Mbo I Alw I Mbo I Hpg II Hpg III GGAGGACCGAAGGAGCTAACCGCTTTTTTGCACACACAGGATCATGTAACTCGCCTTGATCGTTGGAACCGGAGCT CCTCCTGGCTTCCTCCAAAAAACGTGTTGTACCCCCTAGTACATTGACCGCAACTAGCACCTTGGAACGTTGGCTTGGTACATTGGCTTGGTACATTGGCTTGGTACATTGGCTTGGTACATTGGCTTGGGAACGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAG	CA CA	

	59/68		7N
3200	3280	3360	Fig.
MSP II ScrF I Nci I MSP II Nci I MSP II AVG II HHO I AVG II HHO II AVG I	HPh MSP I MSP I HPA NIG IV NIG IV NIG IV NIG IV SCCGACCGACCAAATAACGACTATTTAGACTCGGCCA GCCGACCGACCAAATAACGACTATTTAGACTCGGCCA 3243	3210 96 II I 1V Mnl I CCAGATGGTAAGCCCTCCCGTATCGTATCTACACGACG GGTCTACCATTCGGGAGGGCATAGCATCAATAGATGTGTGTG	3282

.

Mn	1		
With I	344 40	22	9
Mn I I	CITTAGATTGAT GAAATCTAACTA 3.	MSe I HSe I CCTTAACGTG GGAATTGCAC 3512	Bstu I Hinp I Hha I in GCGCTAATC
Mn! I I	TTACTCATATAT AATGAGTATATA	NIG III BSPH I ICATGACCAAAAT AGTACTGGTTTTA 1496 3496	Squ3A I Mbo I Dpn I A I w I s 4 Y I
Mn! I TGCCTCACTGATTAAGCATTGGTAA ACGGAGTGACTAATTCGTAACCATT ACGGAGTGACTAATTCGTAACCATT TGCCTCACTGATTAAGCCATTGGTAA ACGGAGTGACTAATTCGTAACCATT S338 S338 S338 S338 S338 S338 S338 S480 S5003 S480 S480 S478 S478 S478 S478 S478 S478 S478 S480 S460 S478 S460 S467 S478 S478 S467 S478 S478 S478 S467 S478 S478 S478 S467 S478 S478	II TGTCAGACCAA ACAGTCTGGTT	SA I I I CTTTTTGATAAT GAAAACTATTA	Sau3A I Mbo I Dpn I Mbo II BstY I A W I GGAICTICTIG
Mn I Ms TGCCTCACTGATT ACGGAGTGACTAA ACGGAGTGACTAA 75 3379 Mse I Mbo Mse I Bppn Mse I Bppn ATTTAAAAGGAT TAAATTTTCCTAA TAAATTTTCCTAA 3466 3466 3466 3466	e I MGE AAGCATTGGTAA TTCGTAACCATT 88 3398	Squ Mbo II Mbo II Hph I	Sau3A Wbo I Dpn I AGAAAGAICA
	Mn I IV I GCTCACTGAT GGGAGTGACTA S 3379	Sau Mbo Mbo Dra I Bbh I Alw I A ATTTCCTA TAAATTTCCTA 1 A ATTTCCTA 3459 3466 3459 3466 3466 3466	ga I CGTCAGACCCG

61/68				
	3680	3760 I I	3840 Fig.	
3590 3590 3591	I ACCAACTCTTTTTCCGAAG TGGTTGAGAAAAAGGCTTC	Hde III de I GGCCACCACTTCAAGAACTC CCGGTGGTGAAGTTCTTGAG 1740 740 3741 Scr F No i Msp Hpd	CGATAAGTCGTGTCTTA GCTATTCAGCACAGAAT	
3575 3576 3576 3576 3576	du 3A I bo I W I I I I I A I U ATCAAGAGCT TAGTTCTCGA 3658	651 CCGTAGTTA GGCATCAAT GGCATCAAT	CTGCCAGTGG GACGGTCACC	
3564 3564 3564 3565 3565	3555 SS MSP HPG T+h111 II TGTTTGCCGC ACAACGCCG ACAACGGCCG 3648 3648	MDB I TCTAGTGTA AGATCACATA 3723 3723 Fnu4 AIWN I B	CCAGTGGCTG GGTCACCGAG	
3557 3557 3557	NspB II ACCAGCGGTGGTT TGGTCGCCACCAA 3631	CAAATACTGTC GTTTATGACAC	GCTAATCCTGTTA CGATTAGGACAAT	
3533	Tthill II Bby I Tthill II GCTGCTTGCAACAAAAAACCACCGCT GCTGCTTGCAACAAAAAAAACCACGGCT 3601 3609	Mae III GTAACTGGCTTCAGCAGAGCGCAGATAC CATTGACCGAAGTCGTCTCGCGTCTATG 3699	Mn! I TGTAGCACCGCCTACATACCTCGCTCT ACATCGTGGCGGATGTATGGAGCGAGA	

	62	2/68		79
	3920	4000	4080	Fig.
$\begin{array}{cccccccccccccccccccccccccccccccccccc$	NSPB II FRU4H I Bbv I Bb	HinP Hha I Hae II CGTGAGCATTGAGAAAGCGCC GCACTCGTAACTCTTTCGCGG 3973 3973		

WO 98/289/	L		
		63/68	7R
	4160	4240	1 4320 Fig.
4069	GGGCGGAGCCTATGGAAA CCCCGCCTCGGATACCTTT 4147	II 4 I 1 CTTTCCTGCGTTATCCCC AGAAAGGACGCAATAGGGG	T XY 1
4056 4056 4056 4056	I CTCGTCAGG GAGCAGTCC	NSPH I NSPH I NSP752 Afl II CCTCACATGT CGAGTGTACA CGAGTGTACA 4216 4216 4216 4216 4216	Fnu4H I Bbv I nu4H I cccccc ccccccc 6ccTcccc 4294 4294
4046	I SfaM TTTTTGTGATG AAAACACTAC 4129	Hae III Hae I CTGGCCTTTT GACCGGAAAA 4203 4203	GATACCGCTCGCTATGCCGAGCCTATGCCGAGCCAGCCAG
4040	Mni I Hga I ACCTCTGACTTGAGCGTCGATGGAGACTGAACTCGCAGCTAGAGTGAACTCGCAGCTAGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGGTGGAGGA	Hae III ScrF I EcoR II BstN I NIG IV 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	Alu TTACCGCCTTTGAGTGAGCT AATGGCGGAAACTCACTCGA 4277
4011 4018	ATAGTCCTGTCGGGTTTCGCC	Hae II] BstU I BstU I	Hinf I IGATTCTGTGGATAACCGTA ACTAAGACACCTATTGGCAT 4242

ϵ	64/68
4400	1 I I I I I I I I I I I I I I I I I I I
### HinP I HINP	A-A-A

ig. 7T

4560 Hpa II

66/68

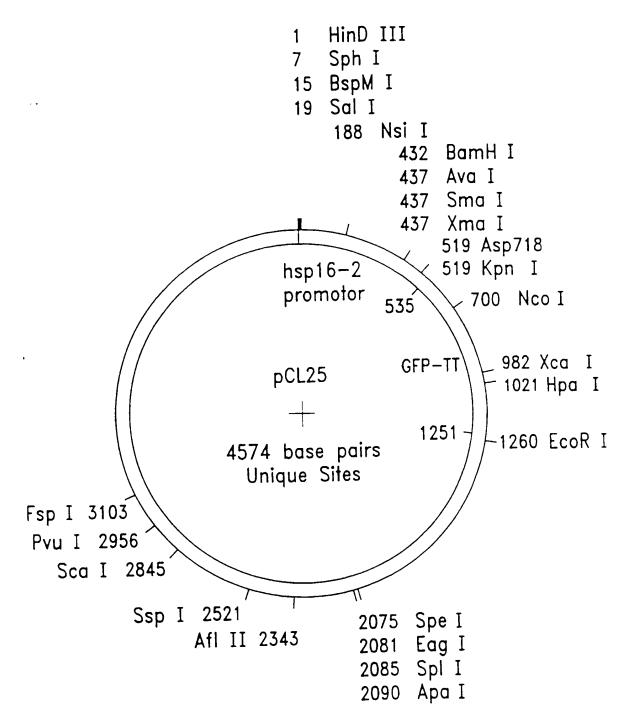


Fig. 8

atcacaatga catcettatg tatatgcatt ctttttattg ttteetgata tttattetet

67/68

aaggagagga aggaccagat gggcacgatg gacaagacgg agtcccagga ggaccaagag gaccaccagg agaggctggg agagacggaa acccaggaat gccaggagac aggatecgat ggaaageeag gateteeagg aggaaaggga ggatgaaaaa cacttggacc acaaggacca aatggagcte caggacteag aggaatgege aggaaaagat gecgaatact geaaatgeee aggaagagaa ggagatgetg caggacagic iggaccacaa gagaaattcc cagatggtat accaaatgga gaggaaatca atgtcaatge actgttgaga actcatgeee accaggaeea caaagtteee aggaggtgga tteecagatg gteeatteee gaatggagga atggagecae tggtgtteag ecaecageae caacteeaaa eecatatgga gacttettgt aaatataaaa attteagaee agaaggaeea eeaggageee aagatgeega agatgtteaa aaeaeeeeae eaaeaggatg etteaeetgt tggttaaact tggagcagga accgcttcca accgtgtgag acgtcaacaa agattcatga cattttcca atatetttag atetaaetga aaattteeag atgaeeetaa etaeggegae gteeggegee atatetggaa tgagetggat geegaaateg eeaaetteag agtgagttte tttctaatic acigataati tcaattatti agagteteae tgaagaeatg aaaattttt gttttgtatt tggtttacca tggatgttaa etttateatt teaettttte ttggeaeteg tecaaaaate teaattaaaa atttetggat gccttcaagg atatggagga gctgctggag aggacgggta attgtatttt etggagecae tttgttggtt tetetitttg eegetgette getttaeagt cacagggacc aactggaaga gatgettate egeaaattee aattgtagae attaaaatac aaaactgtat gggccgcccd acgtccagga caagccagcc agctccacca gacaaccagg caggggcccc agagaccatt cagacgtcat catactictc ggggaaaata ggagagcag aaaaataaaa ccagaaggac gacgatggag gacgaagtgc ccacaaggac caggacticc ggaaatcaac aatattaaac ggagctcgfg tgtggaccac ggaccacgtg gctggaccag ggatatggag Hgacggaa ctatatagca atggaggat gaactttett caagtttcca gggttgata 261 541 1201 141 361 421 081 841 781 901 481 661 721 961 241 301 601

Fig. 9

68/68

DCGPPGAPGSDGKPGSPGGKGDDGERPLGRPGPRGPPGEAGPEGPQGPTGRDAYPGQSG KFPGGGFPDGPFPNGGGPRGGNQCQCTVENSCPPGPAGPEGEEGPDGHDGQDGVPGFD GKDAEDVQNTPPTGCFTCPQGPLGPQGPNGAPGLRGMRGARGQPGRPGRDGNPGMPG **LGAGTASNRVRRQQYGGYGATGVQPPAPTPNPYGGYGASQPAPPEKFPDGIPNGGNQP** PQGEPGLQGYGGAAGEDGPEGPPGAPGLPGKDAEYCKCPGREGDAGRSARRHRKFQL MTLTTATSGAIVFSGATLLVSLFAAASLYSQVSNIWNELDAEIANFRSLTEDMWVDMVK

Fig. 10



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6: C12N 15/00, A01K 67/033, C07K 14/435, 14/47

(11) International Publication Number:

WO 98/28971

(43) International Publication Date:

9 July 1998 (09.07.98)

(21) International Application Number:

PCT/US97/23819

A3

(22) International Filing Date:

31 December 1997 (31.12.97)

(30) Priority Data:

08/778,645 60/080,222

US 3 January 1997 (03.01.97)

US 3 January 1997 (03.01.97)

(71) Applicant (for all designated States except US): UNIVER-SITY TECHNOLOGY CORPORATION [US/US]; Suite 250, 3101 Iris Avenue, Boulder, CO 80301 (US).

(72) Inventor; and

(75) Inventor/Applicant (for US only): LINK, Christopher [US/US]: 1837 Spruce Street, Boulder, CO 80302 (US).

(74) Agents: MONROY, Gladys, H. et al.; Morrison & Foerster LLP, 755 Page Mill Road, Palo Alto, CA 94304-1018 (US). (81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR. BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, GW, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX. NO. NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU. MC, NL, PT. SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published

With international search report.

Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(88) Date of publication of the international search report:

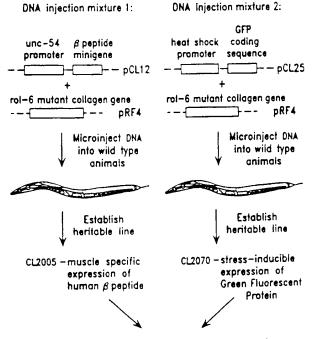
29 October 1998 (29.10.98)

(54) Title: TRANSGENIC ANIMALS FOR IN VIVO VISUALIZATION OF BETA-AMYLOID TOXICITY

(57) Abstract

personal and

The present invention relates to methods and compositions for visualization of the toxic effects of transgenes in vivo. In particular, the present invention provides methods and compositions for the production and use of transgenic, including dually transgenic, Caenorhabditis elegans for visualization of the toxic effects of B-amyloid accumulation in VIVO



Mate lines together, recover dual transgenic animals by classic genetic techniques

PCT/US 97/23819

A CLASSIFICATION OF SUBJECT MATTER C07K14/47 CO7K14/435 IPC 6 C12N15/00 A01K67/033 According to International Patent Classification (IPC) or to both national classification and IPC B FIELDS SEARCHED Minimum documentation searched (classification system followed by classification symbols) A01K C07K IPC 6 Excumentation searched other than minimum documentation to the extent that such documents are included in the fields searched Electronic data base consulted during the international search (name of data base and where practical search terms used) C. DOCUMENTS CONSIDERED TO BE RELEVANT Relevant to plain No Citation of document, with indication, where appropriate of the relevant passages CA 2 088 379 A (CANDIDO EDWARD P M 1-4.6.7.χ 13-15.20 :STRINGHAM EVE G (CA): JONES DONALD (CA)) 30 July 1994 1 - 23see the whole document Υ 1-8. LINK, C.D.: "Expression of human 10-15. beta-amyloid peptide in transgenic 18-20. Caenorhabditis elegans" 22.23 PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF USA.. vol. 92, no. 20, 26 September 1995, pages 9368-9372, XP002065639 WASHINGTON US cited in the application 11-15 see the whole document χ -/--Patent family members are listed in annex Further documents are listed in the continuation of box C X Special categories of cited documents T" later document published after the international tiling date or priority date and not in conflict with the application but A document defining the general state of the art which is not considered to be of particular relevance. cited to understand the principle or theory underlying the E" earlier document but published on or after the international "X" document of particular relevance, the claimed inventio cannot be considered novel or cannot be considered to tiling date involve an inventive step when the document is taken alone document which may throw doubts on pnority claim(s) or which is cited to establish the publication date of another Y" document of particular relevance, the claimed invention cannot be considered to involve an inventive stap when the document is combined with one or more other such document. citation or other special reason (as specified) Of document referring to an oral disclosure use exhibition or ments, such combination being obvious to a person skilled other means document published prior to the international filing date but document member of the same patent family later than the priority date claimed Date of mailing of the international search repor Date of the actual completion of theinternational search 1 1 09 1998 2 September 1998 Authorized officer Name and mailing address of the ISA European Patent Office, P.B. 5818 Patentiaan 2 NL - 2280 HV Rijswijk Tel (+31-70) 340-2040 Tx 31 651 epo ni Chambonnet, F Fax (+31-70) 340-3016

4

Int. Itional Application No PCT/US 97/23819

Continu	ation) DOCUMENTS CONSIDERED TO BE RELEVANT	
189.10	Citation of document, with indication where appropriate, of the relevant passages	Relevant to down No.
	LINK, C.D.: "Transgenic Caenorhabditis elegans as model system to study amyloid formation and toxicity" NEUROBIOLOGY OF AGING. vol. 15, no. suppl. 1, 29 July 1994, page S33-XP002065640 see the whole document	1.5-7
X	WO 95 11968 A (ATHENA NEUROSCIENCES INC :LILLY CO ELI (US)) 4 May 1995 see page 38. line 3 - page 39. line 22	11.12.19
Y	WO 93 23533 A (TSI CORP) 25 November 1993 see page 27, paragraph 3 - page 28, paragraph 2	11-15
Y	CHALFIE M ET AL: "GREEN FLUORESCENT PROTEIN AS A MARKER FOR GENE EXPRESSION" SCIENCE. vol. 263. no. 5148. ll February 1994. pages 802-805. XP002003599 see the whole document	1.2.6.7. 9.11. 13-17. 20.21
Y	MOSS, J. B. ET AL.: "Green fluorescent protein marks skeletal muscle in murine cell lines and zebrafish" GENE. vol. 173, no. 1. 1996. pages 89-98. XP002065641 AMSTERDAM NL see the whole document	1.2.9. 13.16.17
	•	

Int. Ational application No PCT/US 97/23819

Boxi	Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)
This inte	emational Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons
1	Claims Nos because they relate to subject matter not required to be searched by this Authority, namely
	··
2.	Claims Nos . because they relate to parts of the international Application that do not comply with the prescribed requirements to such an extent that no meaningful international Search can be carried out, specifically
3	Claims Nos because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6 4(a)
Box II	Observations where unity of invention is lacking (Continuation of Item 2 of first sheet)
This In	nternational Searching Authority found multiple inventions in this international application, as follows
	see additional sheet
1. [3	As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims
2.	As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee
3	As only some of the required additional search fees were timely paid by the applicant, this international Search Report covers only those claims for which fees were paid, specifically claims Nos.
4 [No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims, it is covered by claims Nos.
Ren	The additional search fees were accompanied by the applicant's protest X No protest accompanied the payment of additional search fees

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

and partially 11-23 1. Claims: 1-10

a transparent animal comprising:

(a) a toxic transgene;

- (b) an inducible promotor operably linked to a reporter gene; wherein the toxic transgene induces the promotor and wherein the expression of the reporter gene is detectable in a method for providing such a dually transgenic transparent non-human animal comprising :
 - a) providing:
 - i) a first and a second non-human transparent animals;

ii) a first toxic transgene; and

- iii) a second transgene, comprising an inducible promotor operably linked to a reporter gene, wherein the toxic transgene induces the promotor and wherein the expression of the reporter gene is detectable in vivo,
- b) introducing said first toxic transgene into the genome of said first non-human animal to produce a first transgenic animal and introducing said second transgene into the genome of said second non-human animal to produce a second transgenic animal; and
- c) mating said first transgenic animal with said second transgenic animal to produce a dually transgenic transparent non-human animal wherein said toxic transgene is expressed and induces the promotor so that the said reporter gene is also expressed and detectable in vivo;
- 2. Claims: partially 11-23

a method to produce a dually transgenic non-human animal with a toxic transgene and a reporter are expressed comprising :

a) providing:

i) a first and a second non-human animal;

ii) a first toxic transgene; and

- iii) a second transgene, comprising a reporter,
- b) introducing said *irst toxic transgene into the genome of said first non-human animal to produce a first transgenic animal and introducing said second transgene into the genome of said second non-human animal to produce a second transgenic animal; and

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

c) mating said first transgenic animal with said second transgenic animal to produce a dually transgenic non-human animal wherein said toxic transgene and said reporter are expressed

in the cases non relevant of the first described subject, that means where it is not a transparent animal comprising :

(a) a toxic transgene;

(b) an inducible promotor operably linked to a reporter gene; wherein the toxic transgene induces the promotor and wherein the expression of the reporter gene is detectable in vivo.

information on patent family members

Ir ational Application No PCT/US 97/23819

Patent document cited in search report		Publication date	Patent family member(s)	Publication date
CA 2088379	Α	30-07-1994	NONE	
wo 9511968	A	04-05-1995	US 5604102 A US 5612486 A AU 8079894 A AU 8080994 A CA 2174429 A CA 2174632 A EP 0736106 A EP 0730643 A JP 9508196 T JP 9507746 T WO 9511994 A	18-02-1997 18-03-1997 22-05-1995 22-05-1995 04-05-1995 04-05-1995 09-10-1996 11-09-1996 19-08-1997 12-08-1997 04-05-1995
WO 9323533	 A	25-11-1993	AU 4240293 A	13-12-1993